UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

The influence of a schizophrenia polygenic risk score on white matter microstructure in schizophrenia, bipolar disorder and health

Beatriz Candeias Simões

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil de sinais e imagens médicas

Dissertação orientada por:
Dr. Diana Maria Pinto Prata
Prof. Hugo Alexandre Ferreira

2018

Acknowledgements

This thesis would not have been possible without the guidance and help of many people. At first, I would like to thank my supervisors, Dr. Diana Pinto Prata and Dr. Hugo Alexandre Ferreira for helping me overcame my struggles. Both of my supervisors were attentive and arranged meetings every time I had a question related to my project. Due to their availability, I was able to conclude my project and learn new tools and methods.

Furthermore, I want to thank all my colleagues from dprata lab for being present during my thesis period and for supporting me. Their intervention during the presentations I did in the lab meetings was very helpful and contributed for my growth as a master student. I am especially grateful to Vânia Tavares, Mónica Costa and Katja Brodman that guided me through the last year and helped me to find balance during this period.

A special thanks to King's Collage London for providing the data that I used in my project. Without their contribuition it would have been impossible to conduct this project. I am grateful for the help that Emma Mallas and Mariana Balseiro gave me in solving questions about the methodology of this study. Moreover, I am thankful for the help that Inês Almeida, Marta Silva and Gonçalo Cosmo gave me while writing my thesis in LaTeX.

I am thankful for the help and support that my friends Marta Silva, Sara Ferreira, João Peixoto, Diogo Cunha and Inês Almeida gave me. They listened and advised me, so I could overcome my personal and professional battles.

At last, I want to thank all my family that provided me with all I needed to conclude my master's in Biomedical and Biophysics Engineering.

Abstract

Schizophrenia (SZ) and bipolar disorder (BD) are classified as psychiatric disorders by the Diagnostic and Statistical Manual of Mental Health Disorders fifth edition (DSM-5). Both illnesses are highly heritable, share symptomatology, and have a polygenic architecture. Given the polygenic architecture of these disorders, several genetic risk variants can be combined in a compound polygenic risk score (PRS) previously associated with the illness, and its impact can be tested on brain phenotypes that have also been associated with the illness, in a so-called neuroimaging genetics approach. To further validate and better characterize a recently reported PRS for SZ, I have assessed its impact on white matter (WM) microstructure, which is known to be impaired in SZ, and to a lesser extent, in SZ relatives and in BD. Each participant of this study was genotyped in order for its SZ PRS to be calculated. The PRS used was compounded using the genetic risk variants found in the latest Psychiatric Genomics Consortium (PGC) schizophrenia meta-analysis study, and was found to explain almost 10% of the variance in SZ status. In the present study, I estimated the effect of this SZ PRS on brain WM microstructure, using two proxy features [fractional anisotropy (FA) and mean diffusivity (MD)], previously collected with diffusion tensor imaging (DTI) in a Caucasian group of SZ (n=22), BD (n=25), relatives (REL; n=28) and healthy controls (HC; n=68). The DTI images of each participant were pre-processed and analysed with FSL software. Resultant of the pre-processing steps, FA and MD images were created. Using tract-based spatial statistics (TBSS) approach, the FA and MD images were transformed in 4D volumes (*skeletons*) in which the fourth dimension represents the subject ID. The final step of the analysis consisted in feeding the 4D volumes into a voxel-wise cross-subject general linear model, with PRS and diagnostic group as independent variables and FA and MD as dependent variables. Main effects and interaction effects were inferred. No significant correlation between PRS and FA or MD was found, either as a main effect or a diagnosis-dependent effect. Nevertheless, small main effect trends revealed a positive association of PRS with FA in uncinate fasciculus, corpus callosum, middle cerebellar peduncle and anterior thalamic radiation. In addition, other main effect trends of PRS revealed a negative association with FA in the superior longitudinal fasciculus, posterior and anterior thalamic radiation, corticopontine tract, middle cerebellar peduncle and cingulum. Regarding MD, positive main effect trends of PRS on this measure were found in the inferior longitudinal fasciculus and posterior corona radiata, while negative effects were found in inferior cerebellar peduncle. There were also no significant diagnosis by PRS interactions on FA, albeit, trend-level effects showed that PRS and FA have higher correlation in REL than SZ patients in the middle cerebellar peduncle. Main effects of diagnosis on FA and MD are reported for completeness. In sum, I have found no evidence to support a role for the current SZ PRS genetic risk factors in contributing for the known decreased FA and increased MD in SZ and BD. However, given the small penetrance of SZ genetic genetic variants, and the small sample size of this group, this approach should be replicated in a larger sample size, before an effect of PRS on WM microstructure can be excluded.

Keywords: schizophrenia, bipolar disorder, polygenic risk score, diffusion tensor imaging, fractional anisotropy.

Resumo

Doenças psicóticas como a esquizofrenia e a doença bipolar afetam a cognição e o comportamento de um indivíduo e são altamente hereditárias. Dado que estas doenças partilham entre si semelhanças e diferenças, estudos têm focado os seus esforços para desvendar o efeito das variações genéticas nos diferentes fenótipos. Diversas variações genéticas têm sido associadas à esquizofrenia e à doença bipolar, no entanto cada uma destas variações apenas explica uma porção muito pequena da doença. Contudo, estas variações genéticas podem ser combinadas num *score* que explica uma maior porção da variância entre doentes e sujeitos saudáveis. A esta *score* dá-se o nome de *score* de risco poligenético (PRS). Neste estudo, o valor de PRS foi calculado para cada indivíduo com base em variações genéticas associadas com esquizofrenia extraídas da mais recente meta-análise do *Psychiatric Genomics Consortium* (PGC).

Além de variações genéticas, alterações estruturais no cérebro têm sido reportadas em estudos de neuroimagiologia. Especificamente, estudos têm associado alterações na integridade da substância branca com esquizofrenia e a doença bipolar. Estas alterações estão normalmente associadas a processos biológicos como redução do grau de mielinização e variação da permeabilidade das membrana axonais a água. Para estudar a integridade da substância branca são utilizadas técnicas de imagem como imagiologia por tensor de difusão (DTI). Com esta modalidade de ressonância magnética é possível extrair medidas de difusão como anisotropia fraccional (FA) e a difusividade média (MD). Dado que a esquizofrenia e a doença bipolar estão ambas associadas ao PRS e a alterações da integridade da substância branca, a hipótese deste estudo é que, por consequência, o PRS está associado a alterações na substância branca. Além disso, estudos anteriores revelaram que uma redução nos valores de FA e um aumento nos valores de MD estão associados a esquizofrenia e doença bipolar. Por isso, espera-se que o PRS esteja negativamente associado com a FA e positivamente associado com MD, sendo que se espera um efeito mais forte para a esquizofrenia e a doença bipolar.

Neste estudo, 22 esquizofrénicos, 25 doentes bipolares, 28 familiares saudáveis e 68 participantes saudáveis foram incluídos na análise. A análise das variáveis demográficas na amostra evidenciou diferenças significativas entre grupos de diagnóstico quanto à idade. Cada participante foi genotipado para que o seu PRS fosse calculado. Este *score* foi determinado como a soma dos alelos de risco que cada sujeito possui ponderado pelo logaritmo de *odds ratio*. Posteriormente, cada imagem foi pré-processada usando FSL. O pré-processamento consistiu (1) na correção de cada imagem de DTI quanto a correntes de Foucault e ao movimento, (2) na extração dos vóxeis relativos ao cérebro e (3) no ajuste de um tensor de difusão a cada voxel da imagem. Como resultado do pré-processamento, imagens de FA e MD foram criadas para cada participante. Para fazer uma análise de grupo foi usado *tract-based spatial statistics* (TBSS), um método que investiga todo o cérebro e estima um esqueleto para FA/MD em toda a amostra. Este esqueleto é posteriormente incluído numa análise entre participantes através de todos os vóxeis usando um modelo linear generalizado. Neste estudo foram realizados 4 tipos de análises: (1) a análise da influência das variáveis demográficas na FA/MD, (2) a análise do efeito dos diferentes grupos de diagnóstico na FA/MD, (3) a correlação entre o PRS e a FA/MD e (4) a análise dos efeitos de interação entre PRS e o grupo de diagnostico na FA/MD.

Para além da amostra referida, a análise foi também realizada para apenas uma porção desta amostra, controlada para as diferenças de idade entre grupos. Nesta amostra foram incluídos 18 doentes com esquizofrenia, 25 doentes bipolares, 25 parentes saudáveis e 47 sujeitos saudáveis sem relação de parentesco com os doentes.

Os resultados deste estudo revelaram que a idade, o quociente de inteligência (QI) e os anos de educação de cada sujeito influenciam significativamente a FA na corona radiata anterior e no fascículo uncinado, respetivamente. Sendo que a idade está negativamente associada à FA e as restantes variáveis estão positivamente associadas. O género e o QI mostraram ter um efeito significativo na MD no pedúnculo cerebelar e corona radiata anterior, respetivamente. Quanto ao diagnóstico, os grupos de doença mostraram estar associados a valores mais baixos de FA quando comparados com os grupos saudáveis. Efeitos significativos foram encontrados no corpo caloso. Para uma amostra controlada para as diferenças de idade, os grupos doentes mostraram também ter valores de FA reduzidos em comparação com os grupos saudáveis em vóxeis localizados na corona radiata posterior e trato corticospinal. A influência do diagnóstico na MD mostrou ter um comportamento antagónico em relação aos resultados encontrados para a FA na mesma análise. O grupo dos doentes bipolares teve um efeito maior na MD que os grupos saudáveis, no entanto este efeito não atingiu a significância. Para a análise dos efeitos da PRS na FA nenhum efeito foi significativo. Para além disto, os resultados não foram conclusivos quanto à direção da influência do PRS na FA, dado que foram encontrados efeitos negativos e positivos. Os efeitos do PRS na MD foram, tal como para a FA, pequenos e não significativos. Na análise dos efeitos de interação entre a PRS e o diagnóstico na FA foram encontrados efeitos em que a correlação entre o PRS e o diagnóstico era maior nos doentes bipolares quando comparados com o grupo de esquizofrenia e os controlos no pedúnculo cerebelar médio. Para além disto, os resultados mostraram também efeitos em que a correlação entre o PRS e o diagnóstico se revelou maior para os familiares em comparação com os esquizofrénicos e os controlos no pedúnculo cerebelar médio e no fascículo fronto-ocipital inferior, respetivamente. No entanto, todos estes efeitos foram obtidos para o nível não corrigido, ou seja, não atingiram significância. Para a MD, apenas um efeito não corrigido foi encontrado no fascículo fronto-ocipital inferior, sendo este efeito estabelecido para o contraste em que a correlação entre o PRS e a MD era superior para controlos em comparação com doentes bipolares.

Com base nos resultados descritos, é possível concluir que a influência de diferentes grupos de diagnóstico na FA/MD é distinta. Como foi mencionado, valores reduzidos de FA estão associados a grupos com doença o que sugere que os indivíduos pertencentes a estes grupos sofrem alterações ao nível da integridade da substância branca, em resultado de provável desmielinização. A MD mostrou ter um comportamento contrário à FA, uma vez que traduz a magnitude de difusão e não a direccionalidade da mesma. A aquisição de ambas as medidas é uma mais valia para a interpretação dos resultados. Relativamente ao PRS, nenhuma relação significativa deste score com as medidas de difusão foi encontrada. A explicação mais plausível para tal é que a dimensão da amostra utilizada neste estudo não é suficientemente grande para detetar pequenas variações genéticas. Contudo, é possível que o PRS, calculado pelo PGC, não esteja relacionado com a FA/MD. Tal pode acontecer porque o PRS calculado não incluí as variantes da esquizofrenia que se relacionam com alterações na integridade da substância branca.

Para que se possa tirar conclusões sobre a relação entre o PRS para esquizofrenia e a FA/MD, considerando ou não grupos de diagnóstico, este estudo deve ser replicado numa amostra com uma dimensão superior. Para além disto, os resultados podem também ser mais evidentes se for calculado um PRS que explique uma maior porção da variação entre esquizofrenia e os sujeitos saudáveis.

Palavras-chave: esquizofrenia, doença bipolar, *score* de risco poligenético, imagiologia por tensor de difusão, anisotropia fraccional.

Contents

List of Figures

List of Tables

List of Abbreviations

ASD	**A**utism **S**pectrum **D**isorder
BACS	**B**rief **A**ssessement of **C**ognition in **S**chizophrenia
BD	**B**ipolar **D**isorder
CPT-IP	**C**ontinuous **P**erformance **T**ask- **I**dentical **P**airs
DANVA	**D**iagnostic **A**nalysis of **N**on-**v**erbal **A**ccuracy
DLPFC	**D**orsal **L**eft **Pref**rontal **C**ortex
DSM	**D**iagnostic **S**tatistical of Manual of **M**ental Health Disorders
DTI	**D**iffusion **T**ensor **I**maging
EY	**E**ducational **A**ttainment
FA	**F**ractional **A**nisotropy
FACT	**F**ibre **A**ssignement by **C**ontinuous **T**racting
FEP	**F**irst **E**pisode **P**sychosis
FIGS	**F**amily **I**nterview for **G**enetics **S**tudies
GLM	**G**eneral **L**inear **M**odel
GWAS	**G**enome-**w**ide **a**ssociation **s**tudies
HC	**H**ealthy **C**ontrols
IoPPN	**I**nstitute **o**f **P**sychiatric, **P**sycology and **N**euroscience
IQ	**I**ntelligence **Q**uotient
MD	**M**ean **D**iffusivity
MDD	**M**ajor **D**epressive **D**isorder
MHT	**M**orey **H**ouse **T**est
MRI	**M**agnet **R**esonance **I**maging
PGC	**P**sychiatric **G**enomics **C**onsortium
PRS	**P**olygenic **R**isk **S**core
PTSD	**P**ost**t**raumatic **S**tress **D**isorder
REL	**Rel**atives
RD	**R**adial **D**iffusion
SNP	**S**ingle **N**ucleotide **P**olymorphism
SZ	**S**chi**z**ophrenia
TBSS	**T**act-**b**ased-**s**patial-**s**tatistics
TEACH	**T**est of **E**veryday **A**tention for **Ch**ildren
TFCE	**T**hreshold **F**ree **C**luster **E**nhacement
TRACULA	**Tra**cts **C**onstrained by **U**nder**l**ying **A**natomy
VBM	**V**oxel-**b**ased-**m**orphometry
VLPFC	**V**entrolateral **L**eft **Pref**rontal **C**ortex
WAIS-III	**W**echsler **A**dult **I**ntelligence **S**cale -III
WAIS-R	**W**echsler **A**dult **I**ntelligence **S**cale -**R**evised
WISC-III	**W**echsler **I**ntelligence **S**cale for **C**hildren -III

WM	**W**hite **M**atter
WRAT	**W**ide-**R**ange **A**chievement **T**est
YE	**Y**ears of **E**ducation

1 Introduction

Psychotic disorders are a group of disorders that affect the mind and influence an individual's capability of understanding reality [1]. Psychotic disorders, including schizophrenia (SZ) and bipolar disorder (BD) with psychotic features, are often chronic, debilitating illnesses that result in lifelong limitations in psychosocial functioning, significant caregiver burden, and substantial economic costs [2].

Both SZ and BD are highly heritable, but their common and specific etiological and pathological causes are not fully understood. Due to this, SZ and BD remain chronic and recurrent for most cases, entailing high rates of morbidity and mortality [2]. Uncovering similarities and differences between these disorders will have important implications for psychiatric nosology, clinical management and treatment, and our understanding of the aetiology of psychotic illnesses [3]. Due to this, a great effort has been made to study the etiological factors of these illnesses, such as genetic risk factors.

Emerging evidence suggests shared genetic susceptibility between SZ and BD [4]. Like most complex disorders, they are likely to occur due to a combination of genetic and environmental factors. Recent developments in genetic research have contributed to improve our understanding of the genetic architecture of these complex disorders. The genetic risk of having a complex mental disorder is conferred by many polymorphisms that regulate different biochemical pathways related with the phenotype. Previous studies have established that SZ is highly polygenic, with many common genetic variants that can lead to abnormal gene expression, consequently contributing to the risk of developing this disease. In a genome-wide association studies (GWAs), 108 independent regions were associated with SZ [5]. It is very important to improve risk prediction of developing complex disorders such as the ones referred above. However, each genetic marker individually only explains a small portion of the genetic variation, which has insignificant predictive power. Another approach relies on examining disorder prediction by summarizing variation across all the associated genetic loci into a quantitative score, like the generation of polygenic risk scores (PRSs). The calculation of this score has been successfully associated with SZ [6].

Neuroimaging studies have reported the effect of some polymorphisms, such as one of the ZNF804A gene in white matter (WM) microstructure integrity in SZ and BD [7]. Fractional anisotropy (FA) and mean diffusivity (MD) are proxies for the WM microstructure integrity and can be extracted from images acquired using diffusion tensor imaging (DTI). FA describes the directionality of the diffusion of water molecules in the axons, while MD represents the magnitude of the diffusion [8].

The goal of this project is to assess the effect of the SZ PRS (based on the previously associated 108 regions) on WM microstructure integrity, using as a proxy, FA and MD. To accomplish this, I tested for a

relationship between FA and SZ PRS. Considering that both a high PRS and a low FA are associated with SZ and BD, I hypothesized that (1) FA values were negatively associated with SZ PRS (i.e FA decreases with an increase in PRS), (2) The effects of SZ PRS on FA/MD were different between the different diagnostic groups. Moreover, I hypothesized that MD had FA's opposite behaviour on both PRS and diagnosis.

1.1 Theoretical Background

1.1.1 Psychotic and Mood Disorders

Psychotic disorders are a group of disorders that are defined as abnormalities in one or more of the five following domains: delusions, hallucinations, disorganized thinking (speech), grossly disorganized or abnormal motor behaviour, and negative symptoms [1].

In more detail, delusions are fixed beliefs that are not amenable to change in light of conflicting evidence. Hallucinations are perception-like experiences that occur without an external stimulus and that can occur in any sensory modality. Disorganized thinking is typically inferred from the individual's speech and reveals disconnection on the semantic context. Grossly disorganized or abnormal behaviour manifests itself in a variety of ways, such as catatonic behaviour [1]. On the other hand, negative symptoms, such as anhedonia (inability to experience pleasure) and apathy (state of indifference), reveal a reduction in normal functions, particularly in emotions and behaviours.

The group of psychotic disorders includes SZ, schizophreniform disorder, brief psychotic disorder, schizoaffective disorder, delusional disorders, schizotypal disorder and catatonia. The differences between these disorders rely on the level of psychosis, the duration of the illness and the number of symptoms [1].

On the other hand, BD and related disorders are characterised by symptoms as maniac and hypomanic episodes and depressive episodes. A maniac episode is a distinct period of abnormally and persistently elevated, expansive or irritable mood, lasting at least one week. A hypomanic episode only differs from a manic episode in duration and severity, in this case, lasting at least 4 consecutive days and present most of the day, nearly every day and with less severity. In contrast, a depressive episode is characterised by symptoms as depressed mood, diminished interest or pleasure in almost all activities, insomnia, fatigue, psychomotor agitation, feeling of worthlessness or guilt, difficulties in concentration and recurrent thoughts of death [1].

BD is considered by the DSM-5 as a bridge between psychotic and depressive disorders in terms of symptomatology, family history, and genetics. Due to the similarities between BD and SZ, authors have proposed a model that suggests a continuum spectrum (Figure 1.1) between two extremes, SZ and BD [4][9].

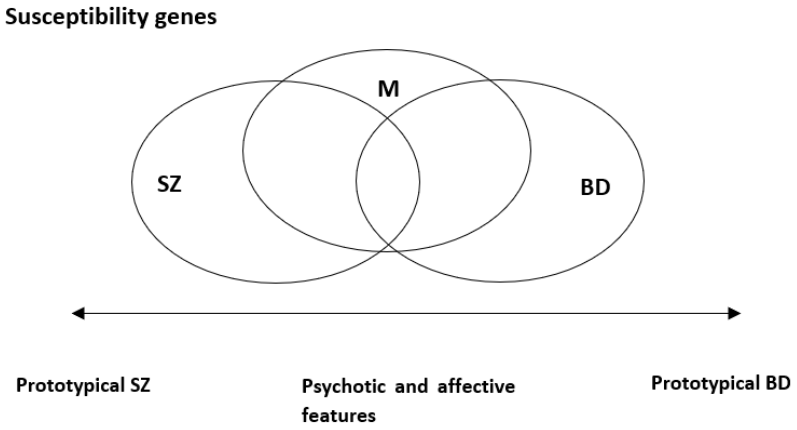

FIGURE 1.1: The relationship between the clinical disorders in the psychosis-bipolar spectrum and susceptibility genes specific to schizophrenia (SZ), bipolar disorder (BD) and schizoaffective disorder, SZ and BD (M). Adapted from [4].

1.1.1.1 Schizophrenia

SZ is among the most common of the severe mental illnesses, it is a chronic psychotic disorder with a lifetime prevalence of about 1% of the world population [10]. As referred before, SZ falls within the scope of psychotic illnesses indicating the presence of symptoms that will lead to misinterpretation of reality. Patients with SZ can suffer from three types of symptoms: positive symptoms (hallucinations, delusions and thought disorganization), negative symptoms (anhedonia and social withdrawal) and cognitive symptoms (impaired working memory and executive function).

The DSM-5 diagnostic criteria for SZ requires at least the presence of two of the five characteristic symptoms of psychotic disorders, and these symptoms must be present for a minimum period of a month, as a part of a total illness duration off a minimum of six months. SZ is diagnosed only if one of the first three symptoms (delusions, hallucinations and disorganized thinking) is present [1].

Although the abnormalities in brain development associated with SZ may begin in utero, childhood-onset SZ is relatively uncommon with an incidence of less than 0.04% [11]. Typically, the onset of SZ occurs in the late adolescence or early adulthood, with a slightly later onset in females [12].

1.1.1.2 Bipolar Disorder

BD is a recurrent chronic disorder characterised by fluctuations in mood state and energy, normally, between mania or hypomania and depressive episodes. This disorder affects more than 1% of the world's population despite of nationality, ethnic origin and socioeconomic status [13].

The criteria for the diagnosis, by DSM-5, of bipolar I disorder (BD-I) is the presence of a maniac episode. The maniac episode may have been preceded or followed by hypomanic or depressive episodes. All these symptoms have been explained in the beginning of these chapter.

About two-thirds of BD patients have at least one psychotic symptom during manic episodes [14]. Bipolar II disorder (BD-II) is characterised by one major depressive episode and at least one hypomanic episode. A major depressive episode may include depressed mood, loss of interest or pleasure in almost all activities, weight loss, insomnia, hypersomnia, fatigue, reduced ability to think and thoughts of death [1].

1.1.2 Genetics

Genes are fundamental to our understanding of the mechanisms involved in diseased states, particularly, complex disorders, such as SZ and BD [15].

Emerging evidence suggests that SZ and BD share genetic susceptibility [4] and like most complex disorders, they are likely to occur due to a combination of genetic and environmental factors. As referred before, the genetic risk of having a complex mental disorder is conferred by many polymorphisms that regulate different biochemical pathways related with the phenotype, being SZ highly polygenic. GWAS are studies in which the goal is to find associations between genomic variants and disorders. Genomic variants are alterations in the most common DNA nucleotide sequence and include single nucleotide polymorphisms (SNPs), variations in a single nucleotide that occurs in a specific position in the genome [16].

The reasoning behind the GWAS is that if specific variants are found more frequently in patients than in their HC, then the variants may be indicative of a genetic association with the disease [17]. In a GWAS, 108 independent regions were associated with SZ [5]. GWAS were also conducted for BD and these studies report significant evidences of association for CACNA1C and identified a new intronic variant ODZ4. A combined GWAS analysis for BD and SZ yielded strong association evidence for SNPs in CACNA1C [18]. However, studying each genetic marker individually only explains a small portion of the genetic variation, which has insignificant predictive power [19]. An alternative approach is the generation of PRSs [20].

1.1.2.1 Polygenetic risk score

As referred before, a PRS is a number based on the variation across multiple genetic loci. A PRS is calculated for each subject as a weighted sum of the known risk variants they carry, which are estimated using the results of independent large GWAS. Each variant is weighted by the logarithm of the odds ratio [6]. The genetic variants used in the calculation of the PRS are the ones below a pre-specified p-value threshold. For higher p-value thresholds, the number of variants or SNPs included in the PRS decreases [6], [21].

Studies that used PRS methodology revealed that a PRS for SZ, explains a considerable fraction of the variance between SZ and HC. Moreover, PRS explains a substantial fraction of the hereditability of SZ, with considerable similarities with BD [22]. The most recent study has collected 108 variants that explain the polygenetic nature of SZ, increasing the proportion of variance in this illness [5].

1.1.2.2 Imaging Genetics

Imaging genetics refers to the use of anatomical or physiological imaging technologies as phenotypic assays to evaluate the influence of genetic variations.

There are three basic principles for imaging genetics (1) selection of genetic variants, (2) the control for non-genotype factors, as environmental influences, and (3) the selection of appropriate methodology.

The variants or candidate genes selected should have known neurobiological consequences in order to emphasize a specific mechanism through which genes have an impact on brain structure and related behaviour. It Is important to control for non-genotype factors while performing imaging genetics analysis, because potential genetic effects are still relatively small when compared with larger effects such as age, sex and environmental influences. For functional neuroimaging, the relatively small genetic effects also demand a well characterised behavioural task. This task must be both sensitive and specific to the brain mechanism under investigation [15].

Several functional and structural studies have focused on SZ and BD and their intermediate phenotypes. An intermediate phenotype is a quantitative biological trait that is reliable and reasonably heritable, ie, shows greater prevalence in unaffected relatives (REL) of patients than in the general population. An example of an intermediate phenotype is WM volume [23]. Typically, the approaches used in imaging genetic studies are GWAS or PRS analysis.

GWAS have identified significant risk genes for SZ (such as ZNF804A, CACNA1C, TCF4, NCAN and NRGN) and BD (ZNF804A, CACNA1C, PBRM1, ANK3, NCAN and DGKH) [2] [24] [25]. The human ZNF804A gene codes for a protein that contains one C2H2 type zinc-finger domain suggesting a role in DNA binding and transcription [7]. CACNA1C is a gene that encodes an alpha-1 subunit of a voltage-dependent calcium channel. This subunit forms the pore through which ions pass into the cell [26]. TFC4 is a protein coding gene for the transcription factor TFIIIC subunit TFC4, and NRGN or neurogranin encodes a postsynaptic protein kinase substrate that binds calmodulin in the absence of calcium [27]. ANK3 is the protein coding gene for AnkyrinG that is involved in neuronal development and signalling [28]. DGKH encodes a member of the diacylglycerol kinase enzyme family, but the pathophysiological role of this enzyme is still elusive [29]. NCAN encodes neurocan, a brain-specific chondroitin sulfate proteoglycan that is thought to influence neuronal adhesion and migration [30]. At last, PBRM1 encodes a subunit of ATP-dependent chromatin-remodelling complexes. Mutations at this locus have been associated with primary clear cell renal cell carcinoma [31].

However, GWAS do not explain how these variants contribute to the complex disorders previously mentioned. Due to this, the approach imaging genetics is fundamental to show how risk genes impact

brain function and structure [2]. Previous literature has reported neuroimaging phenotypes that are affected by risk genes [2], these are presented in Table 1.1. The findings revealed by PRS analysis are going to be referred in the literature review.

TABLE 1.1: Neuroimaging phenotypes grouped in two major categories: structural phenotypes and functional phenotypes, which have been found to be affected by common variations in genes implicated in SZ and/or BD risk.

Neuroimaging phenotypes					
Structural phenotypes					
white-matter	volume	density	grey-matter	cortical folding	thickness
ANK3 ZNF804A	CACNA1C ZNF804A TCF4	ZNF804A	CACNA1C NRGN TCF4 ZNF804A	NCAN	ZNF804A
Functional phenotypes					
regional activation in executive tasks		functional connectivity in executive tasks		facial affect recognition	theory-of-mind
ANK3 CACNA1C DGKH ZNF804A		CACNA1C ZNF804A		CACNA1C ZNF804A	ZNF804A

1.1.3 Imaging white matter microstructure

The WM microstructure characterises and quantifies suspected myelin pathology. These alterations in myelination of WM pathways have been associated with impairments in brain connectivity resulting in diseased conditions. Quantitative information about WM microstructure is crucial to identify the aetiology of a disease [32].

Advances in magnetic resonance imaging (MRI) techniques have made possible to measure putative proxies for WM microstructure, as FA and MD, using DTI [33].

1.1.3.1 MRI basics

MRI is an imaging modality that creates a map of the hydrogen nuclei in different tissues. This map results from the signal produced when nuclei alter they state of energy in the presence of an external magnetic field. The image intensity depends on the number of protons in any spatial location, as well as physical properties of the tissue such as viscosity, stiffness and protein content. In comparison with other imaging techniques, MRI has the advantage of using a magnetic field instead of ionizing radiation [34].

MRI generates spatial images with high resolution enabling the diagnosis of several types of pathologies and, therefore, is widely used in the clinical context [35].

1.1.3.2 Diffusion Tensor Imaging

DTI was originally proposed by Peter Basser in 1994 [36], [37]. The introduction of this technique allowed, for the first time, a rotationally invariant description of the shape of water diffusion. The invariance to rotation enabled the application of DTI to the complex anatomy of the fibre tracts. DTI has been used in a variety of studies including SZ, traumatic brain injury, autism and aging [8].

DTI is a technique that measures the diffusion of water molecules in biological tissues, including fibres tracts. The diffusion of water molecules in fibre tracts is not the same in all directions (isotropic), it does vary with direction (anisotropic). The anisotropic diffusion of water molecules in tissue, like WM, is due to tissue heterogeneity, in this case, caused by cellular membranes and myelinization of axons. This anisotropy can indicate the underlying tissue orientation [8].

The diffusion tensor describes the diffusion of water molecules using a Gaussian model (Equation 1.1). In Equation 1.2, Δt represents the time interval, Δr the displacement and D the diffusion tensor that is a 3×3 covariance matrix (Equation 1.2).

$$P(\Delta r, \Delta t) = \frac{1}{\sqrt{(4\pi\Delta t)^3 |D|}} exp\left\{\frac{-\Delta r^T D^{-1} \Delta r}{4\Delta t}\right\} \tag{1.1}$$

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{1.2}$$

The diffusion tensor, D, describes the covariance of diffusion displacements in a 3D space normalized by the diffusion time. The diagonal elements are nonzero values that correspond to the diffusion variances along x, y and z axes, the off-diagonal elements are the covariance terms and are symmetric about the diagonal. The diagonalization of the diffusion tensor results in the eigenvalues (λ_1, λ_2, λ_3) and corresponding eigenvectors (ε_1, ε_2, ε_3), which describe the directions and apparent diffusivities along the axes of principal diffusion [38]. Together, the eigenvalues and the eigenvectors define an ellipsoid that represents the isosurface of diffusion probability (Figure 1.2). Isotropic diffusion is characterised by nearly equal eigenvalues. In the other hand, anisotropic diffusion is characterised by significantly different eigenvalues [8].

The diffusion tensor is a probe for characterizing both normal and abnormal tissue microstructure, being the abnormal tissue related with diseased states.

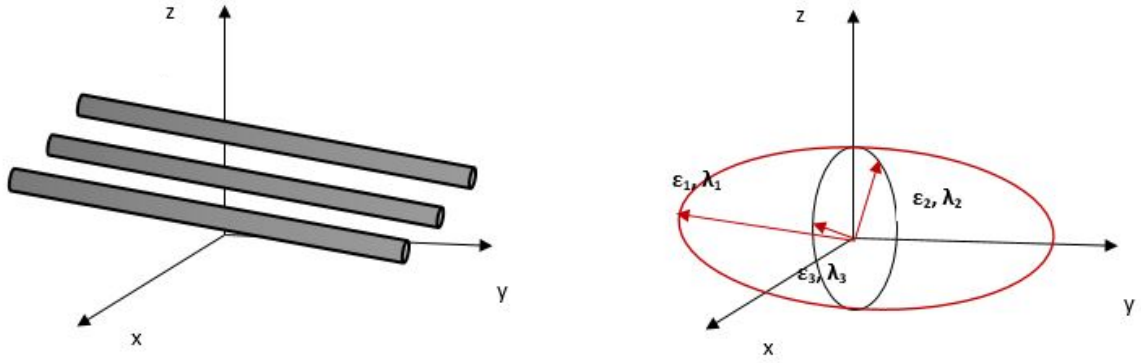

FIGURE 1.2: Schematic representations of diffusion displacement distributions for the diffusion tensor (right) and its relation with axons that are represented by tubes (right). Ellipsoids are used to represent diffusion displacements. λ_1, λ_2, λ_3 represent the eigenvalues and ε_1, ε_2, ε_3 the eigenvectors. Adapted from [8].

To measure diffusion using MRI, magnetic field gradients are employed to create an image that is sensitive to diffusion in a particular direction. The effect of these gradients "cancels out" for stationary water molecules and causes a random phase shift for molecules that diffuse. Due to this, the signal from diffusing molecules will be lost creating darker voxels. This means that WM tracts will appear dark in the diffusion weighted image for that direction [8].

The Stejskal-Tanner equation (Equation 1.3) describes how the signal intensity at each voxel decreases in the presence of Gaussian diffusion. In this equation, S represents the diffusion weighted signal, S_0 is the diffusion weighted signal without any diffusion gradients, D is the diffusion coefficient, and b is the LeBihan's factor (Equation 1.4) [38]. The b-factor describes the pulse sequence, gradient strength and physical constants. Higher b-factors correspond to stronger diffusion effects. In the Equation 4, G represents the magnitude of the gradient pulse, γ the proton gyromagnetic ratio, δ the duration and Δ the separated time interval (Figure 1.3) [39]. The "optimal" choice of b-factor is not clearly defined and depends upon field strength, number of signals averaged, anatomical features, and predicted pathology.

$$S = S_0 e^{-bD} \tag{1.3}$$

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \tag{1.4}$$

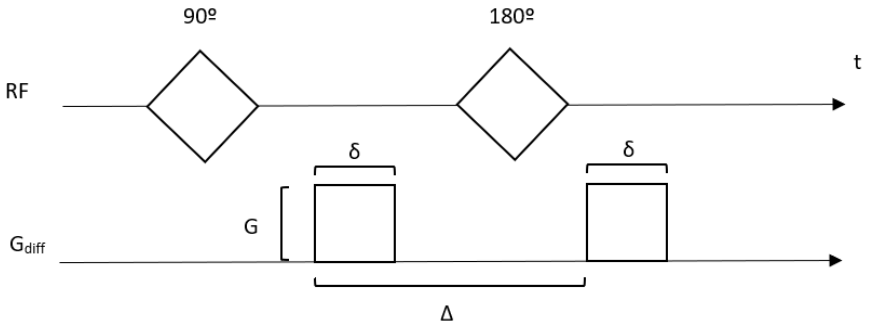

FIGURE 1.3: Schematic of a diffusion weighted pulse sequence. G represents the magnitude of the pulse, δ the duration, Δ the separated time interval and RF the radiofrequency pulse and the diffusion gradient (G_{diff}) . Adapted from [39].

The two most common measures used in DTI are MD and FA. MD is a scalar that results from the average of the tensor's eigenvalues. A similar quantity to the MD is the trace of the tensor corresponding to the sum of the eigenvalues. Both measures are related to the total amount of diffusion in a voxel, which is related to the amount of water in the extracellular space [8].

As referred before, the most widely used measure of anisotropy is FA described by Equation 1.5, where λ_1, λ_2, λ_3 are the tensor eigenvalues.

Values of FA can range from 0 to 1, being zero purely isotropic (diffusion being equal in all directions) and one purely anisotropic (diffusion only along a specific direction). FA appears to be sensitive to a broad spectrum of pathological conditions. However, this measure cannot describe the full tensor shape and distribution. Measurements of the MD may help to better understand how the diffusion tensor is changing [38].

$$FA = \sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{2\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2\right)}} \tag{1.5}$$

1.2 Literature Review

1.2.1 Schizophrenia polygenic risk score studies

As referred before, SZ is highly heritable and the study of the phenotypic manifestations of increased SZ liability is extremely important to better understand the disorder.

Several studies have used SZ PRS to establish associations in diverse categories such as (1) psychiatric disorders, (2) symptoms/severity, (3) cognition, (4) brain structure and (5) brain function. A summary of the results of the SZ PRS studies is presented in Table 1.2a and 1.2b.

1.2.1.1 Psychiatric Disorders

A study conducted in the University of Oslo, reported that the SZ-PRS was higher in patients with SZ, BD-I and psychosis when compared with HC [40]. However, another study showed no evidences of an association between SZ PRS and psychosis history in individuals with BD [41].

In yet another study, results have suggested an association between SZ PRS and BD, depression, and autism spectrum disorder (ASD) [42]. Strong evidences were reported for the relationship between SZ PRS and depression, for both severe atypical and typical depression [43]. SZ PRS was also associated with BD in two other studies with different discovery samples [22], [44].

In the Marine Resilience studies, no association between post-traumatic stress (PTSD) and SZ PRS was found [45].

Based on the studies referred before it is possible to detect some consistent evidences regarding the association of SZ PRS in individuals with BD and depression. These evidences are consistent with the genetic overlap between psychiatric disorders [4].

1.2.1.2 Schizophrenia severity/symptoms

Regarding symptomatology, SZ PRS has been associated with different symptom dimensions: positive, negative, mania, depression and disorganization [46]. Another study, reported strong evidences of an association of SZ PRS with the negative symptoms, but not with positive symptoms [47].

A recent study has reported that a higher SZ PRS is associated with adolescent negative symptoms and anxiety disorders, but not with depression or psychotic experiences [48]. A previous study also reported an association of SZ PRS with a decrease in negative symptoms in adolescence [49].

Based on the studies referenced in this subsection it can be understood that there is lack of evidences that support the relation of SZ PRS with positive symptoms in SZ. This could suggest that the SNPs used to calculate the PRS for SZ are more strongly related to other symptom dimension, such as negative

symptoms. However, evidence for an association between SZ PRS and negative symptoms was inconsistent, with studies reporting an association of SZ-PRS with both increased and decreased negative symptoms.

1.2.1.3 Cognition

A higher value of SZ PRS was associated with lower total intelligence quotient (IQ) in a SZ case-control sample and with worse cognitive ability in two different samples [50], [51]. Furthermore, higher SZ PRS was also associated with a greater decline in IQ with age [52].

In another study, there was no evidence of an association between SZ PRS and non-verbal IQ, and weak evidence of association with shorter sustained attention and poorer spatial working memory accuracy [53].Furthermore, a study done by Harvard Medical School and Broad institute of MIT and Harvard studied the effect of the brief assessment of cognition in SZ (BACS), a measure of premorbid intelligence that is minimally affected by psychosis onset, on SZ PRS. It was found that BACS was negatively associated with the values of SZ PRS. Supporting that BACS score is greatly reduced in psychotic disorder patients.

Based on the studies referred before, it is possible to find a consistent association of SZ PRS with poorer cognition. However, there were some inconsistencies regarding the aspects that are more strongly related with SZ genetic risk [54].

1.2.1.4 Brain Function

The use of SZ-PRS in functional imaging studies is an innovative approach that permits the identification of individuals at risk of developing a psychiatric disorder. The first functional study using PRS reported a positive association between SZ PRS with the neural activity in the left dorsolateral prefrontal cortex (DLPFC) and left pars triangularis of the inferior frontal gyrus, while SZ patients and HC preformed a working memory task. However, this study found no significant effects of SZ PRS on diagnosis [55]. A study from the same group, using the same task, confirmed the positive association between SZ PRS and the activity in DLPFC and in the left ventrolateral prefrontal cortex (VLPFC) and additionally in the left frontal medial cortex, including anterior cingulate cortex, in all participants. In comparison with the previous study, this one used a PRS that comprised more variants (600 SNPs) [56].

In a study from the Norwegian Centre for Mental Disorders Research, a negative association between SZ PRS and activation in several brain areas was found, while subjects performed a working memory task. These areas include the inferior frontal gyrus, the middle and superior prefrontal cortex and the right middle temporal gyrus, in both SZ and control groups. This effect was found in a contrast between high and low memory load. Also with this contrast was found that an increasing in SZ PRS is associated with a decrease in activation in the anterior cingulate cortex, right inferior frontal gyrus and insula, and in the bilateral postcentral gyrus [57].

In another study, using a probabilistic learning paradigm, a negative association between SZ PRS and the activation in the right frontal lobe was found, on a whole brain analysis. Moreover, an association was found between the SZ PRS and the activation in the ventral striatum, on a region-of- interest analysis [58]. A study developed in The Centre for Neuroimaging Cognitive Genomics (National University of Ireland), reported that an increasing PRS, calculated using variants on downstream pathway of MRI137 (region in chromosome 1 highly associated with SZ) in HC was associated with increasing activation of a cluster on the right inferior occipital gyrus and right middle temporal gyrus and other in the medial parietal region, during a working memory task [59]. A recent study using a working memory paradigm also, revealed that higher SZ PRS is associated with hypoactivity of the dorsal prefrontal cortex during the manipulation but not maintenance of information, in healthy young adults [60].

Considering the studies referred in this section, it is possible to report some consistent associations between SZ PRS and brain function. For functional studies that focused on working memory there is a consistent association between SZ PRS and the activation of the prefrontal cortex, in patients with SZ. However, the direction of the association is not consistent across studies.

1.2.1.5 Brain Structure

Besides functional studies, SZ PRS has also been associated with brain structure phenotypes in structural neuroimaging studies.

A study conducted in Rudolf Magnus Institute of Neuroscience in University Medical Centre Utrecht, has reported a significant association between SZ-PRS and total brain volume and WM volume, in both SZ patients and HC. From the set of 2020 SNPs used in this study only a group of 186 SNPs showed most evidence for association with WM volume [61].

A study conducted in two different Universities, Frankfurt and Cardiff, used two different subject samples, one with SZ patients, BD patients, REL and HC and another sample with only HC. For the first sample, was found a significant association between SZ-PRS and WM volume. For the sample comprising only HC it was found a negative association between SZ-PRS and WM volume, controlling for age and diagnosis [62].

Another study reported a negative association between SZ PRS and hippocampal volumes, in a sample with patients with first episode psychosis (FEP) and at risk mental state. Moreover, a higher SZ PRS was significantly associated with a higher probability of an individual being assigned to the FEP group in comparison to the at risk mental state group [63]. In a study that used a PRS based on 91 SZ risk variants, no significant effect was found between SZ-PRS and any neuroimaging phenotypes, in a sample of HC [64].

In a recent study with a sample comprising HC and SZ and BD patients, was reported a significant main effect of SZ PRS in left and right hemispheres. It was also found that an increase in SZ PRS is associated with a thinner cortex in SZ and BD patients, but not in HC [65]. Based on the studies mentioned in this section, it is possible to find some consistent associations between SZ-PRS and structural

phenotypes. There are two studies that report an association between SZ PRS and WM volume, in a sample of SZ patients.

1.2.2 DTI studies for SZ and BD

Several DTI studies have reported WM deficits in BD and SZ independently. A summary of the results of DTI studies for SZ and BD is presented in Table 1.3.

Recently, several investigators have focused their studies on the relation between impairments in the WM microstructure and the diagnosis of SZ. In a study with a sample of 94 SZ patients and 91 HC, significant intergroup differences were found. SZ patients demonstrated reduced FA and increased MD compared to HC in the corpus callosum, cingulum, internal capsule, fornix and widespread superficial WM in the frontal, parietal, occipital and temporal lobes. However, the distribution patterns of WM regions with reduced FA and increased MD were similar [66].

A study related SZ symptom dimensions with the deficits in WM microstructure. In two out of five dimensions, i.e. abnormal psychomotor behaviour and negative symptoms, the severity of symptom was specifically associated with WM brain structure. It was also found that patients had reduced FA in bilateral clusters within the corpus callosum, anterior limb of the internal capsule, corona radiata, anterior and posterior thalamic radiation, corticospinal tract, as well as in the superior and inferior longitudinal fasciculi [67].

A study conducted in Indiana University School of Medicine, found that patients had lower FA and higher radial diffusion (RD) than HC in numerous WM tracts, including the corpus callosum and the superior longitudinal fasciculus, in a sample with SZ patients and HC. Illness duration was associated with lower FA, most prominently on the corpus callosum [68].

In a study including 104 SZ patients and 200 HC matched for age, reduced FA values were found through the brain with two types of analysis: voxel-based-morphometry (VBM) and tract-based-spatial-statistics (TBSS). The study also focused on the relationship of FA with age, which differed between patients and HC, with HC showing the gentle FA decline widely noted but patients showing an essentially flat relationship: younger patients had lower FA than HC, but the difference disappeared with age, suggesting that WM deficits in SZ don't progress with age. MD was also measured and was widely increased in patients [69].

A study from the University of California focused on examining the relation between WM microstructure in SZ using DTI and cognitive improvements induced by cognitive training. In comparison with HC, individuals with SZ showed reduced WM integrity at baseline (before cognitive training), in bilateral posterior corona radiata, bilateral retro lenticular internal capsules, bilateral posterior thalamic radiation, left anterior corona radiata, left superior longitudinal fasciculus, left sagittal stratum, right cerebral peduncle and the genu and splenium of the corpus callosum. After cognitive training, participants showed significant gains in attention, processing speed, verbal learning, visual learning and executive learning. These improvements were associated with WM integrity in several brain areas [70].

Several research facilities have also focused their research on studying the relationship between WM deficits and the diagnosis of BD. A DTI tractography study with a sample of 18 BD patients and 16 HC, reported significant differences in FA values between patients and HC. FA values for BD patients were lower compared with HC in the thalamic radiation and uncinate fasciculus [71].

In another study, a significant decline on FA was found in BD patients in all major WM tracts, including cortico-cortical association tracts, i.e uncinate, inferior fronto-occipital, inferior longitudinal, and superior longitudinal fasciculi, interhemispheric tracts, and limbic tracts, like the parahippocampal tract [72].

A study conducted in Washington University School of Medicine that included BD patients with psychosis and without psychosis, reported decreased FA in numerous tracts in BD groups compared to HC: bilateral cingulum-cingulate gyrus bundles, corticospinal tracts, and superior longitudinal fasciculi as well as the right hemisphere cingulum-angular bundle. Only left uncinate fasciculus FA differed between psychotic BD and non-psychotic BD groups. No group differences were found for MD [73]. A recent study with a sample comprising BD patients, non-ill first-degree REL and HC, reported significant group differences in FA values among the three groups for the corpus callosum, the dorsal part of the right cingulum bundle, the hippocampal part of the cingulum bundle bilaterally and the uncinate fasciculus. Moreover, FA values were significantly lower in patients than in HC, and first-degree REL showed a similar behaviour but smaller than in patients [74].

A previous study has investigated the influence of both SZ and BD on the white matter microstructure. Moreover, it has studied the differences between these diseased groups on FA. In both patient groups MD and RD showed a significant increase, while FA decreased compared to the healthy group. Changes in diffusion were located, for both diseases, in the fronto-temporal and callosal networks. In more detail, for SZ FA showed a decrease in respect to HC, while MD and RD showed an increase in widespread WM areas, including corpus callosum, corona radiata, longitudinal fasciculus, internal and external capsule, thalamic radiation. For BD, FA, MD and RD showed, significant differences in many areas, including bilateral internal capsule and corpus callosum, bilateral external capsule, corona radiata, internal capsule and longitudinal fasciculus [75].

Taken all together these studies show reduced values of FA for SZ and BD when compared with HC in several brain areas (consistently in the corpus callosum and thalamic radiations). This pattern suggests that mental disorders, as SZ and BD are associated with WM microstructure impairments, which may be one of the causes of the onset of these disorders. The similarity of the results for SZ and BD suggests there may be a relation between WM deficits and the overlap of mechanisms shared by these two illnesses, reinforcing the importance of studying both disorders together. Moreover, the results are consistent with the global disruption in myelination in complex disorders [66], [69], [75].

1.2.3 Polygenic risk score in DTI studies

Previous literature has used PRSs and assessed its effect on WM microstructure in illness and health. In a study with a sample comprising 70 familial risk of mood disorder and 62 controls, it was found that higher PRSs calculated for major depressive disorder (MDD) relatives were associated with a decrease in WM microstructure integrity (i.e lower FA values), with a peak in the right superior longitudinal fasciculus. However, no significant effect was found for the influence of BD PRS and FA values [76].

Another study focused on assessing the association of PRS for SZ, BD and MDD, generated by the Psychiatric Genomics Consortium (PGC), with WM integrity using the UK Biobank (n = 816). This dataset included patients and controls, but no comparison between groups was performed. In more detail, this study only focused on finding the association between PRSs and FA/MD of the whole data set. No significant associations were found between the PRSs calculated for SZ, BD and MDD and FA/MD [77].

In summary, the results found in the above mentioned literature suggest that WM microstructure may not be closely linked to the genetic mechanisms of major psychiatric disorders.

1.3 Objectives and Hypothesis

The major goal of this project is to find if there is an association between a SZ PRS and WM microstructure integrity, thus combining genetics and neuroimaging techniques. In other words, the goal is to discover to what extent a SZ PRS can predict WM microstructure changes, with FA and MD as proxies. Moreover, I want to assess whether this effect is different in HC vs REL (of SZ or BD patients) vs SZ vs BD populations.

The determination of both FA and MD is important to interpret the changes in the measured diffusion tensor, given that FA represents the direction of the diffusion and MD its magnitude [38].

Considering that a high SZ PRS [40], [41], a decreased FA [66]–[75] and increased MD [66], [69], [75] are associated with SZ and BD, my main hypothesis is that SZ PRS will be negatively associated to FA and positively to MD. In addition, I also expect that the effect of SZ PRS on FA/MD is different between the four diagnostic groups.

The findings of this study are expected to contribute to a potential use of SZ PRS as a diagnostic marker to SZ and BD, to validate both the previous association of the PRS polymorphisms with SZ as well as the decreased FA, and increased MD, as an imaging phenotype of SZ and BD, and to further our understanding of the pathophysiological mechanisms leading to the onset of these illnesses and their underlying brain changes.

1.4 Innovative contributions

This work focuses on unravelling the relationship between the risk of developing SZ and the WM microstructure integrity of the brain in different conditions, this being its main contribution. Although the effect of PRS on brain structure/function has been previously investigated using sMRI, fMRI and DTI with samples including SZ and BD patients, this is the first time that SZ PRS is studied with DTI and tests its influence on SZ and BD patients' WM integrity separately. In more detail, this study is the first one to assess the influence of a SZ PRS on WM integrity in SZ, BD, REL (of SZ and BD patients) and of HC separately and to test how different the influence of SZ PRS is the separate diagnosis groups (i.e. to test to for diagnosis by PRS interactions).

Methodologically, to the best of my knowledge, no study has tested the influence of SZ PRS on WM integrity using a permutation approach to determine the clusters (voxles) of significant effect.

1.5 Collaborations

The data used in this work was collected by researchers from the Institute of Psychiatry, Psychology Neurosciences (IoPPN) in King's College London, United Kingdom. The subject recruitment, acquisition of functional and structural images and the assessment of clinical and neurophysiological measures

TABLE 1.2A: List of previous studies with associations between SZ PRS and diagnosis, symptoms, cognition, functional and structural phenotypes. Bipolar disorder (BD), schizophrenia (SZ), posttraumatic stress disorder (PTSD), major depressive disorder (MDD), first-episode-psychosis (FEP), autism spectrum disorders (ASD), IQ (intelligence quotient), WM (white matter), Morey House Test (MHT), Wechsler Adult Intelligence Scale (WAIS -III), Continuous Performance Task-Identical Pairs (CPT-IP), Wechsler Intelligence Scale for children (WISC -III), Test of everyday attention for children (TEACH), Diagnostic analysis of non-verbal accuracy (DANVA), Brief assessment of cognition in schisophrenia(BACS), Wide-range achievment test(WRAT), educational attainment(EY), DLPFC (dorsolateral prefrontal cortex) and VLPFC (ventrolateral prefrontal cortex).

SZ PRS studies			
Psychiatric disorders (diagnosis)			
Author	**Diagnostic groups**		**Phenotype (direction of association)**
Purcell et al. (2009)	BD (955 cases, 1498 controls)		BD (+)
Lancet et al. (2013)	Depressive disorder (9227 cases, 7383 controls); BD (6990 cases, 4820 controls); ASD (4788 trio cases, 4788 pseudo controls, 161 cases, 526 controls)		BD (+), depression (+) and ASD (+)
Tesli et al. (2014)	570 BD, 452 SZ spectrum, 415 controls		BD (+), SZ (+)
Ruderfer et al. (2014)	SZ (9369 cases, 8723 controls); BD (10410 cases, 10700 controls)		-
Stringer et al. (2014)	BD (1998 cases, 1500 controls)		BD (+)
Nievergelt et al. (2015)	PTSD (3494 cases)		PTSD (+)
Milaneschi et al. (2016)	MDD (1530 cases, 1700 controls)		MDD (+)
Symptoms/severity			
Author	**Diagnostic groups**	**Phenotype (direction of association)**	
Derks et al. (2012)	SZ (314 cases, 148 controls)	Positive symptoms (+), negative symptoms (+), mania (+), depression (+) and thought disorganization (+)	
Fanous et al. (2012)	SZ (2454 cases)	negative symptoms (+)	
Sieradzka et al. (2014)	3427 adolescents	negative symptoms (-)	
Jones et al. (2016)	8230 adolescents	negative symptoms (+)	
Cognition			
Author	**Diagnostic groups**	**Neurocognitive assessment**	**Phenotype (direction of association)**
McIntosh et al. (2013)	937 controls	MHT at 11 years old and WAIS - III at 70 years old	IQ (-)
Lencz et al. (2014)	SZ (5446 cases, 5830 controls)	WAIS -III	IQ (-), cognitive ability (-)
Hatzmanolis et al. (2015)	1079 controls	CPT-IP - attention - and N-back task - working memory	non-verbal IQ (-), spatial working memory (-), attention (-)
Hubbard et al. (2016)	14062 children	WISC - III; TEACH; DANVA	IQ (-), cognitive ability (-)
Shafee et al. (2018)	314 psychotic (100 SZ, 143 BD and 71 schizoaffective disorder), 423 controls	BACS; WRAT IV	BACS (-), WRAT(-), EY (-)

TABLE 1.2B: (Continuation) List of previous studies with associations between SZ-PRS and diagnosis, symptoms, cognition, functional and structural phenotypes. Bipolar disorder (BD), schizophrenia (SZ), posttraumatic stress disorder (PTSD), major depressive disorder (MDD), first-episode-psychosis (FEP), autism spectrum disorders (ASD), IQ (intelligence quotient), WM (white matter), Morey House Test (MHT), Wechsler Adult Intelligence Scale (WAIS -III), Continuous Performance Task-Identical Pairs (CPT-IP), Wechsler Intelligence Scale for children (WISC -III), Test of everyday attention for children (TEACH), Diagnostic analysis of non-verbal accuracy (DANVA) ,BACS (brief assessment of cognition in SZ), WRAT (Wide-Range Achievement Test), EY (educational attainment)DLPFC (dorsolateral prefrontal cortex) and VLPFC (ventrolateral prefrontal cortex).

SZ PRS studies			
Brain Function			
Author	**Diagnostic Groups**	**Task**	**Phenotype (direction of association)**
Walton et al. (2013)	SZ (79 cases, 99 controls)	Working Memory	Left DLPFC (+), Left inferior frontal gyrus, pars triangularis (+)
Walton et al. (2014)	SZ (92 cases, 114 controls)	Working Memory	Left DLPFC and left VLPFC (+), Left frontal medial cortex (+)
Kauppi et al. (2015)	SZ (63 cases, 118 controls)	Working Memory	Right inferior frontal gyrus (-), Middle and superior PFC (-), Right middle temporal gyrus (-)
Lancaster et al. (2016)	83 controls	Probabilistic Learning	Right frontal pole (-), Left ventral striatum (+)
Cosgrove et al. (2017)	86 controls	Spatial working memory	Right inferior occipital gyrus and middle temporal gyrus (+), Medial parietal region (+)
	83 controls	Facial processing	-
Brain Strucuture			
Author	**Diagnostic groups**		**Phenotype (direction of association)**
Terwisscha van Scheltinga et al. (2013)	SZ (152 cases, 142 controls)		Total brain volume (-), WM volume (-), grey matter volume (null)
Ortel-Knochel et al. (2015)	24 SZ, 12 REL, 20 BD, 38 controls		WM volume (-)
Harrisberger et al. (2016)	38 At-risk mental state patients, 27 FEP		Hippocampal volume (-)
Voineskos et al. (2016)	107 controls		Frontal and temporal cortical thickness (null)
Neilson et al. (2017)	43 SZ and BD		cortical thickness (-)
	32 controls		-

TABLE 1.3: List of previous studies with associations between DTI measures and SZ/BD. Schizophrenia (SZ), bipolar disorder (BD), relative (REL), fractional anisotropy (FA), mean diffusivity (MD), radial diffusion (RD), tract-based-spatial-statistics (TBSS), voxel-based-morphometry (VBM), fibre assignment by continuous tracting (FACT), tracts constrained by underlying anatomy (TRACULA).

DTI studies for SZ and BD			
Schizophrenia			
Author	**Diagnostic groups**	**Analysis**	**Phenotype (direction of association)**
Zhuo et al. (2016)	SZ (94 cases, 91 controls)	TBSS	FA (-) and MD (+) in the corpus callosum, cingulum, internal capsule and fornix.
Viher et al. (2016)	SZ (40 cases, 40 controls)	TBSS	FA (-) in the corpus callosum, corona radiata, thalamic radiation, corticospinal tract and longitudinal fasciculus.
Hummer et al. (2016)	SZ (40 cases, 21 controls)	TBSS	FA (-) and RD (+) in corpus callosum and longitudinal fasciculus.
Kanaan et al. (2017)	SZ (104 cases, 200 controls)	VBA	FA (-) and MD (+) in corpus callosum
		TBSS	FA (-) and MD (+) in almost every WM tract.
Squarcina et al. (2016)	SZ (19 cases, 35 controls)	TBSS	FA (-), MD (+) and RD (+) in widespread WM areas, including corpus callosum, corona radiata, longitudinal fasciculus, internal and external capsule, thalamic radiation.
Subramaniam et al. (2017)	SZ (48 cases, 28 controls)	TBSS	FA (-) in the bilateral corona radiata, bilateral portion of the internal capsule, thalamic radiations, genu and splenium of the corpus collosum.
Bipolar disorder			
Author	**Diagnostic groups**	**Analaysis**	**Phenotype (direction of association)**
Lin et al. (2011)	BD (18 cases, 16 controls)	FACT	FA (-) in thalamic radiation and uncinate fasciculus.
Ambrosi et al. (2013)	BD (20 cases, 21 controls)	TBSS	FA (-) in uncinate, inferior fronto-occipital,inferior longitudinal, superior longitudinal fasciculi, interhemispheric tracts, and limbic tracts.
Ji et al. (2017)	48 PBD, 24 NBD, 30 controls	TRACULA	FA (-) in patients in the bilateral cingulum-cingulate gyrus bundles, corticospinal tracts, and superior longitudinal fasciculi as well as the right hemisphere cingulum-angular bundle.
Squarcina et al. (2016)	BD (33 cases, 35 controls)	TBSS	FA (-), MD(+) and RD (+) in bilateral internal capsule and corpus callosum, bilateral external capsule, corona radiata , internal capsule and longitudinal fasciculus.
Mahapatra et al. (2017)	16 BD, 15 REL, 16 controls	FACT	FA (-) in patients than controls. Group differences in FA values among the three groups for the corpus callosum, the dorsal part of the right cingulum bundle, the hippocampal part of the cingulum bundle bilaterally and the uncinate fasciculus.

was done by Dr. Cynthia Fu, Dr. Chris Chaddock, Dr. Marco Picchioni, Dr. Sri Kalindindi, Dr. Fergus Kane, Dr. Colm McDonald, Dr. Elvira Bramon and Dr. Timi Toupoulou. Genomic data was extracted and genotyped by Social Genetic and Development Psychiatry laboratory technicians. The processing of the genetic data and calculation of the PRSs was performed by Dr. Evagnelos Vassos.

Finally, the image and statistical analysis were performed by me, under the supervision of Prof. Diana Prata and Prof. Hugo Ferreira (Faculdade de Ciências da Universidade de Lisboa).

2 Methods

2.1 Sample Description

2.1.1 Participants

In this project, 159 subjects with diffusion imaging and PRS data were selected from a data set that contained participants used in previous studies [7] [78]–[83] at the IoPPN, King's College London. From these subjects, 3 were excluded because they were diagnosed as prodromal. Moreover, 2 participants were removed because their ethnicity was not Caucasian. Afterwards, 11 participants were also excluded due to the presence of relatives within the same diagnostic group. The removal of related subjects inside the same diagnostic group was performed to guarantee that all subjects included in the same diagnostic group were genetically independent from each other. The completeness of the demographic data was the criteria to choose which relative should be included in the diagnostic group. If both REL had the same demographic information the choice was random.

The 143 subjects that did not meet the exclusion criteria were divided in four different diagnostic groups:

(a) SZ patients (n=22).

(b) BD-I patients (n= 25), 80% with psychosis.

(c) REL (n=28) of SZ and BD patients, 75% bipolar disorder REL and 25 % schizophrenia REL.

(d) Unrelated HC (n= 68). Participants in the group of unrelated HC were subjects with no previous history of mental illness and no first-degree REL with a psychotic disorder, as assessed using the Family Interview for Genetic Studies.

All participants included in the sample were Caucasian. This is important as the SZ PRS score used has been calculated by [6] for the Caucasian population. The assessment of the diagnosis was performed by a psychiatrist with the criteria of DSM-4 [84]. According to DSM-4, exclusion criteria was applied to participants that had history of significant head injury and recent substance dependency. The sample is heterogeneous in terms of medication containing non-medicated and medicated patients (with different types of treatment).

This project was approved by the National Health Service South East London Research Ethics Committee, UK (Project Mental Health Genetics and Psychosis) reference number 047/04 and all volunteers gave written informed consent at the time of participation.

2.1.2 Demographics

The statistical analysis of the demographic data was conducted using IBM SPSS Statistics 24 [85]. Statistical significance was considered at p-value < 0.05 threshold.

2.1.2.1 Effects of diagnosis on demographic variables and PRS

At first, normality tests (Shapiro-Wilk test) were performed, these tests revealed non-normal distribution of the demographics in at least one of the diagnostic groups. Due to this, to test if there were demographic differences between diagnostic groups non-parametric tests were performed. Kruskal -Wallis test was used for continuous demographic variables and if there were significant differences between groups, post-hoc tests were performed using Man-Whitney U tests. For categorical variables frequency tables were constructed to determine Pearson Chi-Squared (χ^2) values. The statistical analysis was performed for several demographic variables including age, IQ, years of education (YE), gender and handedness.

IQ scores were assessed using four different scales: Wechsler Adult Intelligence Scale – III (WAIS – III) [86], the Wechsler Adult Intelligence Scale – Revised (WAIS – R) [87], the Wechsler Abbreviated Scale of Intelligence – Full Scale IQ [88] or the National Adult Reading Test [89]. To allow the comparison of IQ scores assessed by different tests the values were standardized to z-scores. The z-score was calculated using Equation 2.1, where x represents the IQ score of each participant, μ the average of IQ scores and σ the standard deviation.

$$Z = \frac{x - \mu}{\sigma} \tag{2.1}$$

Statistical tests showed no significant differences between diagnostic groups in terms of gender, handedness and YE. However, there were significant differences in terms of age and IQ (z-scores). Regarding the differences in age, post-hoc analysis revealed that age was significantly different just between BD and HC (U=528.50; p-value=0.005). Concerning IQ (z-scores), post-hoc analysis showed significant differences between SZ and BD (U= 178.50; p=0.04), SZ and REL (U=150.500; p-value=0.002) and REL and HC (U=574.50; p-value=0.02). The results for the analysis of the demographic differences between diagnostic groups are presented in Table 2.1.

To analyse if there were differences in the diagnostic groups in terms of PRS, an analysis of variance (ANOVA) test was performed after confirming that PRS was normally distributed in all groups. The PRS means in each group were significantly different (F=4.687; p-value=0.004). To take conclusions about which groups were significantly different in terms of PRS a multiple comparison test (Tukey's

range test) was performed. Significant differences between SZ and HC groups (mean difference= 0.836; p-value=0.003) in terms of PRS were found.

In Figure 2.1, a box plot with the PRS distributions in the different diagnostic groups is presented. It is possible to notice that the PRS median is bigger for SZ group, followed by REL, BD and HC, as expected. Unexpectedly the PRS median was greater in REL compared with BD, probably because this group contains contribution of both disease groups (SZ and BD). There is only one outlier in the sample, a BD patient (PRS=-1.91). Moreover, PRS from BD and REL show a moderately symmetric distribution, with the median placed on the centre of the box, while SZ and HC present a more asymmetric distribution.

TABLE 2.1: Participants demographic variables per diagnostic group. Age, IQ (z-scores) and YE are presented as mean (standard deviation).

Demographic Variables	SZ (n=22)	BD (n=25)	REL (n= 28)	HC (n=68)	Statistics, df, p-value
Age (years)	39.6 (2.8)	44.0 (2.3)	40.1 (2.3)	35.7 (1.7)	H= 8.975; p=0.030
IQ (z-scores)	-0.47 (0.23)	0.08 (0.25)	0.54 (0.17)	-0.08 (0.1)	H= 14.674; p=0.002
YE (years)	14.8 (0.4)	14.7 (0.7)	15.1 (0.5)	14.6 (0.3)	H=2.753; p=0.431
Gender (M/F)	16/6	10/15	13/15	34/34	χ^2=5.628; p=0.131
Handedness (R/L/M)	(22/0/0)	(24/0/1)	(24/3/0)	(61/5/2)	χ^2=6.283; p=0.392

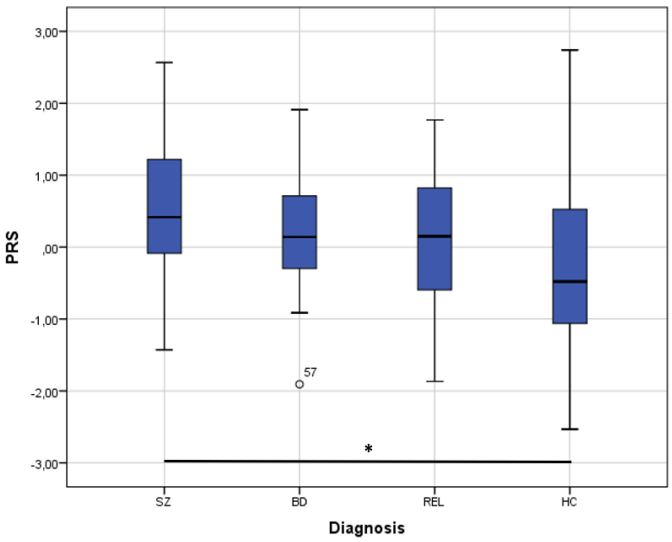

FIGURE 2.1: Box plot of the participants PRS per diagnostic group: schizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC). A significant effect between SZ and HC is showed (mean difference= 0.836; p-value=0.003)

2.1.2.2 Effect of PRS on demographic variables

Spearman's rank correlation coefficient (r_s) was calculated to test if there were significant correlations between PRS and Age/IQ (z-scores)/ YE. The PRS didn't significantly correlate with any of the continuous demographic variables: PRS x Age (r_s=0.013; p-value=0.880), PRS x IQ (r_s= -0.73; p-value=0.389), PRS x YE (r_s= -0.112; p-value=0.181). To assess the relationship between categorical

variables (Gender and Handedness) and PRS a Mann Whitney U test and a Kruskal Wallis test were performed, respectively.

Neither, of these variables was significantly associated with PRS: gender (U=2074.00; p-value=0.124), handedness (H=0.269; p-value=0.874).

2.1.3 Participants subsample

2.1.3.1 Participants

A subsample (n = 115) of the main sample (n = 143) was selected. This subsample included 115 participants that were divided in four groups (SZ, BD, REL and HC) matched for age, i.e with no significant differences in age between diagnostic groups. To select a sample matched for age subjects with ages below 27 were excluded (4 SZ, 2 REL and 21 HC were excluded). This allowed diagnostic groups to have similar age ranges between each other and overcome differences in age between groups.

The rationale behind the selection of this sample relies on the fact that age is strongly related with a decline in the brain WM microstructure integrity, previously reported in [90]–[92]. Across studies, the predominant findings are decreased FA and increased MD as a function of increasing age [93]. Moreover, there are significant differences in age in terms of diagnosis in the sample previously described (n=143). Both these factors increase the risk of having a confounding variable (i.e a variable that influences both the dependent variable and independent variable) in the imaging analysis.

Confounding variables pose a major problem in identifying the real causes of diseases. These variables may erroneously increase or decrease the magnitude of an association, or even invert the direction of these association [94]. Due to this, the control of these variables is extremely important.

The 115 subjects included in the subsample were divided in four groups, as referred before:

(a) SZ patients (n =18).

(b) BD-I patients (n =25), 80% with psychosis.

(c) Healthy REL (n =25) of SZ and BD patients. 76% bipolar REL and 24% SZ REL.

(d) Unrelated HC (n =47). Participants in the group of unrelated HC were subjects with no previous history of mental illness and no first-degree REL with a psychotic disorder, assessed using Family Interview for Genetic Studies, as referred in the previous sections.

Besides age, IQ also shows significant differences between diagnostic groups in the whole-sample (described in the previous section). Moreover, previous studies have reported a significant influence of IQ in WM microstructure [95]. Both the factors stated increase the risk of IQ being a confounder. However, IQ is intimately related with diagnosis as cognitive deficits are related with SZ and BD. Due to this, an IQ-matched sample was not created.

2.1.3.2 Demographics

The statistical analysis of the demographic data was conducted in the same manner as for the whole-sample, described in the previous section.

For the participants subsample there were no significant differences between diagnostic groups in terms of age, gender, YE and handedness. Nevertheless, there were significant differences between diagnostic groups in terms of IQ (z-scores). The results of the statistical analysis of the demographic data are presented in Table 2.2.

The PRSs means in each diagnostic group were significantly different (F=2.748; p-value= 0.046). However, there were no significant differences between groups after multiple comparison tests (Tukey's range test).

In Figure 2.2, a box plot with the PRS distributions in the different diagnostic groups is presented. It is possible to verify that the PRS median is higher in SZ group, followed by REL, BD and HC. The PRS median was greater in REL in compared with BD probably because it contains contribution of both SZ and BD. There are two outliers in the sample, a BD patient (PRS= -0,21) and a HC (PRS= 2,74). Moreover, PRS from SZ, BD and REL show a moderately symmetric distribution, with the median placed on the centre of the box, while HC present a more asymmetric distribution.

The Spearman's rank correlation coefficient was calculated to test if there were significant correlations between PRS and Age/IQ (z-scores)/ YE) and the respective p-values. The PRS didn't significantly correlate with any continuous demographic variable: PRS x Age (r_s=-0.60; p-value=0.521), PRS x IQ (r_s= -0.80; p-value=0.398), PRS x YE (r_s= -0.125; p-value=0.184). The association between PRS and categorical variables was also not significant: Gender (U = 1463; p-value=0.295), Handedness (H = 2.098; p-value =0.350).

TABLE 2.2: Participants subsample demographic variables per diagnostic group for the participants subsample. Age, IQ (z-scores) and YE are presented as mean (standard deviation).

Demographic Variables	SZ (n=18)	BD (n=25)	REL (n= 25)	HC (n=47)	Statistics, df, p-value
Age (years)	43.72 (2.40)	44.04 (2.33)	42.20 (2.34)	41.96 (1.76)	H= 0.940; p=0.816
IQ (z-scores)	-0.37 (0.25)	0.08 (0.25)	0.58 (0.18)	0.02 (0.13)	H= 9.931; p=0.019
YE (years)	14.97 (0.45)	14.68 (0.68)	15.24 (0.55)	14.51 (0.35)	H=3.112; p=0.375
Gender (M/F)	13/5	10/15	13/12	24/23	χ^2=4.408; p=0.221
Handedness (R/L/M)	(18/0/0)	(21/2/1)	(24/0/1)	(44/3/0)	χ^2=6.009; p=0.422

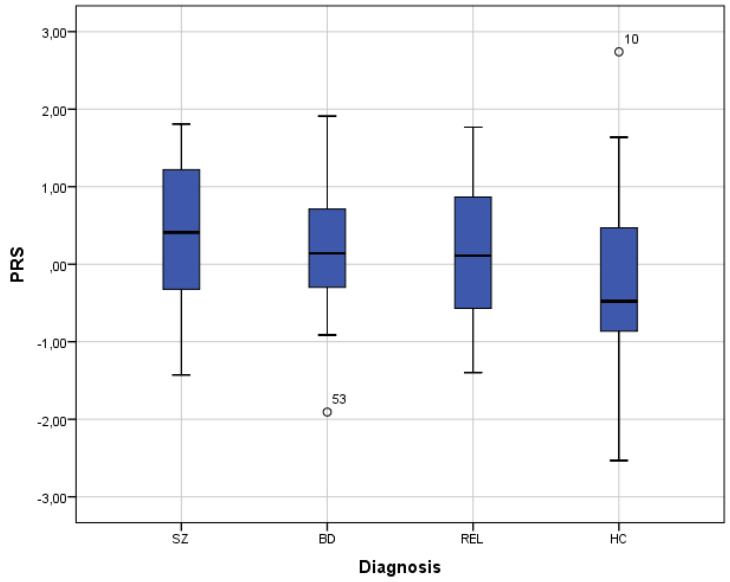

FIGURE 2.2: Box plot of the participants subsample PRS per diagnostic group: schizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC).

2.2 Genetics

2.2.1 Genotyping procedures

DNA was extracted from blood or cheek swabs. The samples were genotyped at the South London and Maudsley NHS Trust/King's College London BRC Genomic Laboratory on the Illumina Human-Core Exome BeadChip. This array provides genetic data for identified genome-wide significant SNPs. Genotypes were processed using GenomeStudio Analysis software version 2011.1 [96].

Quality control included exclusion of SNPs with minor allele frequency < 1%, SNPs and individuals with genotypic failure > 1%, SNPs with Hardy Weinberg equilibrium p < 10-5 in control subjects, mismatch between recorded and genotypic sex, and related individuals. Quality control was performed with PLINK 1.9 [97], [98]. A more detailed explanation can be found in a previous published work from collaborators [6].

2.2.2 Polygenic risk score

The PRS was calculated for each subject, using SNPs found to be significantly associated with SZ on the PGC SZ meta-analysis [5]. The PRSs were computed by a geneticist collaborator at IoPPN using PRSice software [99]. The PGC GWAS identified several risk variants for SZ on a large discovery sample comprising about 36,989 cases and 113,075 controls [5].

The formula for this PRS score has been tested in a target sample included 445 cases with first episode psychosis or SZ, and 265 control subjects, as published previously [6]. Both in this published study's target sample and in the sample used in the present thesis, there is no overlap with the sample in PGC SZ meta-analysis from where SNP were chosen, assuring that there was no inflation of the results. The PRSs were determined for a specific p-value threshold and for each subject (j=1,2,3,...,n) as a sum of the risk alleles of the abovementioned SNP set (i=1,2,3,...,m) that they carried weighted by the logarithm of odds ratio of the discovery sample (Equation 2.2). In Equation 2.2, the term $ln(OR_{SNP(i)})$ represents the weight of each SNP and $X_{SNP(i),j}$ corresponds the SNP genotype for that subject - j - and it can assume the value 0, 1 and 2, depending if the risk allele is absent, in one allele or in both alleles.

$$PRS_{P_T,j} = \sum_{i=1}^{m} ln(OR_{SNP(i)}) X_{SNP(i),j} \tag{2.2}$$

After the calculation of the PRS for each participant a logistic regression was performed to analyse the association between the PRS values and the disease-trait. This was conducted for PRSs with different sets of thresholds (p_T = 0.00000005, 0.00001; 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1). The results of the logistic regression are presented in Table 2.3 .

From Table 2.3, is possible to conclude that for larger thresholds the number of SNPs included in the score increases. The values of PRS included in this study were the ones calculated for the threshold p-value of 0.1, because it assumes the greatest proportion of variance between PRS and the case-control status.

To guarantee that the score comprises informative and independent SNPs, another selection of SNPs was made with minor allele frequency > 10% and imputation information score > 0.9, according to PGC protocol, including only one SNP from the major histocompatibility complex linkage disequilibrium region on chromosome 6 (hg19; chr6:27-33Mb). Principal component analysis was used to model population structure and exclude any outlier individuals using EIGENSTRAT [100]. Ten principal components were used as covariates in the genetic analysis, to control for population stratification.

TABLE 2.3: Results of PRS association with case-control status. It is represented the threshold p-value, the p-value of the association, R^2 and number of SNPs are listed.

Threshold p-value (p_T)	**p-value**	R^2	**Number of SNPs**
0.00000005	0.0067	0.0309	106
0.00001	0.0205	0.0221	353
0.0001	0.0015	0.0419	781
0.001	0.0001	0.0651	1889
0.01	5.78 x 10-5	0.0703	5513
0.05	1.75 x 10-5	0.0802	12479
0.1	**4.20 x 10-6**	**0.0927**	**17965**
0.2	7.10 x 10-6	0.0878	26161
0.5	1.20 x 10-5	0.0829	41866
1	1.77 x 10-5	0.0793	55013

2.3 Imaging

2.3.1 Image acquisition

MRI data was acquired using a 1.5T GE Signal LX system (General Eletric, Milwaukee, WI, USA) in the Mapother House MR unit at the Maudsley Hospital, London, UK, with actively shielded magnetic field gradients (maximum amplitude 40 mT/m). A standard quadrature birdcage head coil was used for both radiofrequency transmission and signal reception.

Diffusion data was acquired using a multi-slice peripherally-gated echo planar imaging sequence, optimized for precise measurement of the diffusion tensor in parenchyma, from 60 contiguous near-axial slice locations for whole brain coverage, with isotropic (2.5 x 2.5 x 2.5 mm) resolution. At each slice location, 7 images were acquired with no diffusion gradients applied (b=0), together with 64 diffusion-weighted images in which gradient directions were uniformly distributed in space.

The pulse sequence parameters were chosen to provide maximum precision in the estimates of the unique elements of the diffusion tensor matrix. These parameters are presented in Table 2.4. Further details are given in a previous article [101].

TABLE 2.4: MRI images acquisition parameters [101].

Parameter	
Echo time (TE)	107 ms
Repetition time (TR)	15 R-R intervals
Duration of the diffusion gradients	17.3 ms
Separation of the diffusion gradients	49 ms
Amplitude of diffusion gradients	40 $mT.m^{-1}$
Maximum diffusion weighted	1300 $s.mm^{-2}$
Field of view	24 cm x 24 cm
Acquisition matrix	96 x 96
Slice thickness/gap	2.5 mm/ 0 mm
Number of slices	60

2.3.2 Image Preprocessing

All images were extracted from **DOUGAL** an image repository from Kings College London. To process the DTI raw data the version 5.0.8 of FSL [102] was used.

Before preprocessing the diffusion images on FSL [95] orientation labels were created using **fslorient** function *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Fslutils)*. Furthermore, to guarantee that all raw images had the same orientation as the standard template image (MNI152) the function **fslreorient2std** *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Orientation%20Explained)* was used. This function applies 90, 180 or 270 - degree rotations on the different axes to get the orientation labels in the same position as in the MNI152 space.

The preprocessing of the diffusion images included three main steps: correction of eddy current induced distortions, brain extraction and tensor fitting (Figure 2.3). These three steps are going to be explained in more detail on the following subsections.

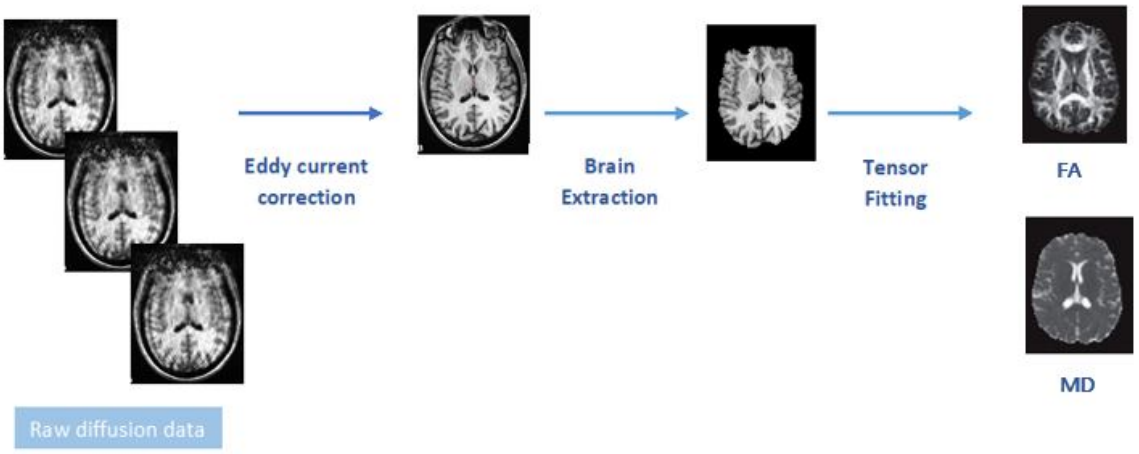

FIGURE 2.3: Representation of the steps involved in the preprocessing of the diffusion imaging data (Eddy current correction, Brain extraction and Tensor fitting).

2.3.2.1 Eddy Current Correction

For diffusion weighted echo planar imaging images the strong, rapidly switched, diffusion encoding gradients are an additional source of off-resonance. The rapidly changing magnetic field induces eddy currents in conductors within the bore, which will in turn induce a magnetic field [103]. To correct eddy current induced distortions in the raw diffusion images the function **eddy_correct** was used *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy)*.

eddy_correct function is an FSL utility that uses FLIRT (FMRIB's Linear Registration Tool) *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT)*, a 12 degree of freedom affine transformation (6 rigid body transformations and 6 of similarity) and correlation ratio as a cost-function to register the diffusion weighted images to a b=0 image (i.e. an image acquired without the application of any diffusion encoding gradient). This model is based on the assumption that two images acquired with the same diffusion gradient will have the same contrast and any differences between them should be related with distortion or/and noise [103].

2.3.2.2 Brain Extraction

To delete non-brain voxels from the diffusion images the function **bet** was used *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET/)*.

This FSL function extracts the brain from the diffusion images by conducting several steps. The first one consists on analysing the image histogram to find robust lower and upper intensity values for the image, and a rough brain/non-brain threshold. The second step focus on finding the centre-of-gravity of the head image along with the thorough size of the head in the image. Finally, the third step consists on conducting a triangular tessellation of a sphere's surface, allowing the slow deformation, one vortex at a time, following forces that keep the surface well-spaced and smooth, while attempting to move toward the brain's edge. If a suitable solution is not arrived the process is rerun with higher smoothness constrains [104].

2.3.2.3 Tensor Fitting

The diffusion tensor has often been used to model local diffusion within a voxel, which can be characterised by several diffusion measures. This model relies on the assumption that local diffusion may be characterised with a 3D Gaussian distribution, whose covariate matrix is proportional to the diffusion tensor [105].

To fit a diffusion tensor model to each voxel of the diffusion images the function **dtifit** was used *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#DTIFIT)*. To use this function a bvals file (file with the different b factors applied in each volume acquisition) and bvecs file (file with the gradient direction at each volume acquisition) are necessary. This function creates different images based on different diffusion measures (including FA and MD) for each one of the subjects included in the sample.

2.3.3 Statistical analysis

To conduct a multi-subject analysis of the diffusion images the TBSS method was used. This method was chosen because it solves alignment and smoothing issues that were present in previous methods, per example VBM [106].

TBSS is a fully automated method that investigates the whole-brain by estimating a *group mean FA skeleton* which represents the centres of all fibre bundles that are generally common to the subjects involved in the study. Each subject's FA data is then projected onto the *mean FA skeleton* in such a way that each *skeleton* voxel takes the FA value from the local centre of the nearest relevant tract. TBSS method can be divided in several steps: nonlinear alignment, identification of the target for alignment, creation of the mean FA and skeleton, projecting individual subject's FA onto the skeleton and statistics and thresholding [106]. These steps are going to be described in more detail on the following subsections.

To perform TBSS on FSL the functions **tbss_1_preproc**, **tbss_2_reg**, **tbss_3 _postreg** and **tbss_4 _prestats** were used (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS/UserGuide). To conduct TBSS in non-FA images the function **tbss_non_FA** was used. Before the nonlinear alignment step the function **tbss_1_preproc** was used to guarantee that all FA images in the working directory were in the right format to perform TBSS.

2.3.3.1 Nonlinear Alignment

To align multiple FA images to each other a nonlinear registration was performed, driven by the FA images themselves. The goal of this step is not to change the fundamental nature of the images, instead the objective is to keep the general structure intact but align the images sufficiently well, so it is possible to conduct the next steps correctly.

The nonlinear registration was based on free-form deformations and B-splines approach. The aim of free-form deformations is to deform an image by moving the control points of an underlying mesh. The warp field applied is found for image positions between the mesh control points using B-spline interpolation. To achieve the optimal warp the control point locations must be moved until the registration cost function is minimised. This cost function attempts to both optimise a voxel-based similarity measure at the same time as imposing regularization (smoothness) on the warp field [106].

To conduct this step on FSL the function **tbss_2_reg** was applied.

2.3.3.2 Identification of the target and creation of the mean FA and skeleton

Upon investigation of the quality of registrations on typical FA images, it was found that registration is more successful if the target is a real FA image rather than an average FA image. Therefore, it was identified a single subject's FA image to act as the target for all nonlinear registrations. The chosen subject should be the most typical subject of the entire group, i.e. the target image that minimises the amount of warping required for all other subjects to align to it. To find this subject every subject was registered to every other subject and the one with minimum distance to all other subjects was chosen [106].

After identifying the most typical subject as the target, all subject's FA images were aligned to it and the entire aligned dataset was affine-transformed into $1 \times 1 \times 1mm^3$ MNI152 space. The transformed FA images were averaged to create the mean FA image. The mean FA image was then fed into a tract skeleton generation, which aims to represent all tracts which are common to all subjects. The skeleton represents each tract as a single line (or surface) running down the centre of the tract [90].

To conduct the steps described in this section the FSL function **tbss_3_postreg** was used.

2.3.3.3 Projecting individual subject's FA onto the skeleton

After creating the mean FA image and it's skeleton every subject aligned FA image was projected onto the mean FA skeleton. The goal of this step is to account for residual misalignments between subjects after the initial nonlinear registrations. To conduct this step each subject's FA image was searched at each point of the skeleton to find the maximum FA value and assign this value to the skeleton voxel. The process described before effectively achieves alignment between the skeleton and the subject's FA images without needing prefect nonlinear preregistration [106].

To perform this step on FSL the function **tbss_4_prestats** was used with a threshold of 0.2. This value is used to threshold the *mean FA skeleton* image to exclude peripherical tracts with significant inter-subject variability and/or partial volume effects due to grey matter.

2.3.3.4 Performing TBSS in non-FA images

To perform TBSS method in non-FA diffusion driven-data the function **tbss_non_FA** was used *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS/UserGuide#Using_non-FA_Images_in_TBSS)*. This function uses FA images to achieve nonlinear registration and skeletonization stages, and to estimate the projection vectors from each individual subject onto the mean FA skeleton. The nonlinear warps and skeleton projection can then be applied to other diffusion images, as MD images.

2.3.3.5 Statistics and Thresholding

After performing the first four steps the data is in the form of a skeletonised 4D image, with the fourth dimension being the subject ID. In this format, the data is ready to be fed into voxelwise multi-subject statistical analysis [106].

To conduct voxelwise multi-subject statistical analysis a general linear model (GLM) is applied. GLMs can be applied to several types of analysis, including multiple regressions, one-sample t-tests, analysis of variance and analysis of covariance. For a single voxel, the GLM can be represented by Equation 2.3.

$$Y = X\beta + \varepsilon \tag{2.3}$$

The variable Y expresses the observed response in terms of a linear combination of explanatory variables in the matrix X plus an error term ε, an independently and identically distributed Gaussian random term. The component X contains the explanatory variables and it is herein called the design matrix. Each column of this matrix corresponds to some effect built into the experiment. The relative contribution of each of these columns to the response, Y, is controlled by the parameter β. These are estimated using standard least squares and inferences about the parameter estimate are made using T or F statistics [107].

To carry out the inferences a permutation-based approach was used. Permutation inferences are powerful while providing excellent control of false positives in a wide range of research scenarios. In permutation methods, the null hypothesis is created simply by observing, after permuting observations, how often the difference between the means would exceed the difference found without permutation [108].

The FSL function **randomise** was used to conduct permutation inferences on a GLM *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise)*. To use this function matrices were created for the experimental

design (X), contrasts (t-tests) and F-tests in a text file and then converted to FSL format using the function **Text2Vest** (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM/CreatingDesignMatricesByHand). The function randomise returned as output test statistic images and threshold free cluster enhacemnet (TFCE)-corrected p-value and uncorrected maps for each statistical test. The TFCE correction aims to enhance areas of signal that exhibit some spatial contiguity without relying on hard-threshold based clustering. The correction output is not intrinsically thresholded, instead this approach serves to improve posterior thresholding. TFCE gives generally better sensitivity than other methods over a wide range of test signal shapes and signal to noise ratio values. The study that explains and introduces this method shows an example of this approach in a real imaging data set, suggesting that TFCE does indeed provide not just improved sensitivity, but richer and more interpretable output than cluster-based thresholding [109].

2.3.3.5.1 Experimental Design To perform these analyses on FSL it was necessary to create several matrices: a design matrix (experimental design), a matrix with contrasts and, depending on the analysis, a matrix with F-tests. For the generation of contrasts, I created contrast vectors comprised of 1s (higher FA/MD), -1s (lower FA/MD) and 0s (hypothesis-free). The F-tests were based on contrasts, so the F-test matrix was only comprised of 1s and 0s used to select which contrasts were going to be included in this statistic test or not, respectively.

The analysis of the effects of diagnosis was performed based on a design matrix that contained the information of the diagnosis for each subject included in the sample. This matrix contained four columns that corresponded to the four groups of diagnosis (SZ, BD, REL and HC) and as many rows as the number of subjects included in the sample. The matrix was coded in 1s or 0s depending on whether the subject was in that diagnosis group or not, respectively. A scheme of the analysis design used to infer the effect of diagnosis on FA/MD is represented in Figure 2.4.

SZ	BD	REL	HC
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
.	.	.	.
.	.	.	.
.	.	.	.
n	n	n	n

FIGURE 2.4: Scheme of the design matrix for the analysis of the effects of diagnosis on FA/MD (without considering any covariates of no interest in the model).

An ANOVA was performed to assess the main effect of diagnosis. To conduct this analysis on FSL three contrast vectors defined as [1 0 0 -1], [0 1 0 -1] and [0 0 1 -1] and the F-statistic was assessed based on these contrasts. These contrasts represent the differences between SZ, BD and REL groups and the

control group. To take conclusions about the direction of the effect, post-hoc analyses were conducted by performing t-tests between the different diagnostic groups.

To determine the effects of PRS on FA/MD a matrix with the 5 columns was created. The first four columns contained information about the group of each subject (as in the design to assess the effects of the diagnosis) and the last column had the values of PRS for each one of the subjects included in the sample (Figure 2.5). This effect was assessed enforcing a linear regression. To conduct this technique on FSL a matrix with contrasts was created containing two contrast vectors defined as [0 0 0 0 1] and [0 0 0 0 -1]. These contrasts represent the negative and positive effects of PRSs on FA/MD, respectively.

SZ	BD	REL	HC	PRS
0	0	0	1	-0.61
1	0	0	0	-1.42
1	0	0	0	0.40
1	0	0	0	-0.44
1	0	0	0	-0.57
0	0	0	1	1.21
0	0	0	1	1.05
0	0	0	1	-0.74
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n	n	n	n	n

FIGURE 2.5: Scheme of the design matrix for the analysis of the PRS effects on FA/MD (without considering any covariates of no interest in the model).

At last, the interaction effects between PRS and diagnosis (PRS x Diagnosis) were inferred. The interaction effects were inferred based on a design matrix composed by eight columns. The first four columns contained the information about the diagnostic group of each subject and the other four columns corresponded to the values of PRS for each diagnostic group (Figure 2.6). For this analysis the PRS x Diagnosis interaction was assessed applying an analysis of covariance (ANCOVA), which allows for the estimation of effects of continuous and of categorical variables and their interaction. For this analysis, t-tests were conducted between the PRSs of each diagnostic group. For example, the contrast vector [0 0 0 0 -1 0 0 1] was implemented to assess the effects in which the correlation between PRS and FA/MD is larger in HC than in SZ patients.

SZ	BD	REL	HC	PRS SZ	PRS BD	PRS REL	PRS HC
0	0	0	1	0	0	0	-0.61
1	0	0	0	-1.42	0	0	0
1	0	0	0	0.40	0	0	0
1	0	0	0	-0.44	0	0	0
1	0	0	0	-0.57	0	0	0
0	0	0	1	0	0	0	1.21
0	0	0	1	0	0	0	1.05
0	0	0	1	0	0	0	-0.74
0	0	0	1	0	0	0	-0.85
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	n	n	n	n	n	n	n

FIGURE 2.6: Scheme of the design matrix for the analysis of the PRS x Diagnosis interaction effects on FA/MD (without considering any covariates of no interest in the model).

The analyses described above were performed with and without demographic variables (age, gender, handedness, YE, IQ) as covariates of no interest. To add a covariate of no interest to the experimental design an extra column was created containing the values of that variable for each subject included in the sample. These variables were added to the model in order to subtract its effect from FA/MD. This was conducted setting the covariate as 0 in the contrast vector.

Besides the main analyses of the study, the analysis of the effects of demographic variables on FA/MD were also assessed using FSL. To assess these effects a column with the values of the demographic variable for each subject were added to the experimental designs. These effects were assessed by setting, in the contrast vector, the position that corresponded to the variable as 1 or -1, representing the negative and positive effects the demographic variable on FA/MD respectively. The effects of the demographic variables on FA/MD are important to create the most optimal experimental design for the analyses of the effects of the variables of interest. For example, if a demographic variable has a strong significant effect on FA, by including it as a covariate of no interest, its contribution to the error term of the model is subtracted, and the model may be more sensitive to detect effects of interest.

After the creation of the experimental design and contrast matrices permutation-inferences on the GLM using **randomise** function were conducted, as described in the previous section.

2.3.3.5.2 Clustering and labelling To extract information about the clusters with p-values below a specific significant threshold the FSL function **cluster** was used *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Cluster)*. This function returned the number of voxels of each cluster, the value of maximum “intensity” inside the cluster, coordinates of the voxel with maximum “intensity” and the location of the centre of gravity within the cluster . For this work, the value of the maximum "intensity" corresponds to the subtraction between one unit and the p-value of the cluster (1- p-value). Statistical significance was achieved for TFCE-corrected results with a p-value < 0.05, while trends were achieved for TFCE-uncorrected results with a p-value < 0.01 [7], [67], [70]. Moreover, clusters were only reported if they overcome the minimum cluster extent of the analysis. For the analysis of the effect of demographics and diagnosis the minimum cluster extent was set to 100 voxels, threshold reported in previous neuroimaging papers [110]–[112]. Due to the few literature regarding the effect of SZ PRS on WM integrity, no minimum

cluster extent was considered for the analysis of the main effect of SZ PRS and for the interaction analysis between SZ PRS and diagnosis (i.e all clusters were reported).

To assess the WM regions/tracts where each cluster was localized the Johns Hopkins University ICBM-DTI-81 white-matter labels atlas was used. If cluster results were retrieved as "No label", labelling was carried out manually using MRI Atlas of Human WM [113]. More information about this atlas is described in Apendix A. Results were overlaid on MNI152 (1mm) standard template and mean FA skeleton.

2.3.3.5.3 Calculation of effect sizes The T-statistic or F-statistic values of the effect's peak were extracted from FA images using the function **fslmeants** *(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Fslutils)*. To help differentiate weaker findings to more robust effects effect sizes were calculated for every reported cluster. Cohen's d effect sizes were computed based upon the uncorrected T-statistic for all effects resulted from a difference between two groups. For the analysis where both dependent and independent variables were continuous (correlations), R^2 effect size was computed for each cluster based on the uncorrected T-statistic. At last, the eta-squared η^2 effect size was calculated for the main effects (ANOVA) based on the uncorrected F-statistic. Cohen's d and R^2 effect sizes were calculated based on Equation 2.4 and 2.5 respectively, where t represents the T-statistic value and DF the degrees of freedom (subtraction of the number of variables and intercept from the number of subjects included in the analysis). η^2 effect size was calculated based on expression 2.6, where F represents the F-statistic, k the number of groups and N the number of subjects.Table 2.5 contains the descriptors for magnitude for Cohen's d, R^2 and η^2 [114], [115].

$$d = \frac{2t}{\sqrt{DF}} \tag{2.4}$$

$$R^2 = \frac{t^2}{t^2 + DF} \tag{2.5}$$

$$\eta^2 = \frac{F(k-1)}{F(k-1) + (N-k)} \tag{2.6}$$

TABLE 2.5: Magnitude of effect size for Cohen's d, R^2 and η^2.

Magnitude of effect	d	R^2	η^2
small	0.2	0.01	0.02
medium	0.5	0.09	0.13
large	0.8	0.25	0.26

3 Results

The results presented in the following subsections are represented in the form of a table, usually followed by a visual representation of the brain regions with significant or trend to statistically significant effects. To better represent some of the statistically significant effects, plots were constructed using IBM SPSS statistics 24. The voxels that survived the significance threshold TFCE-corrected p-value <0.05 or the uncorrected threshold p-value <0.01 and the minimum cluster extent (if applicable), are characterised by the cluster extent (k), the p-value, the 1-(p-value) peak coordinates, the peak T-statistic or F-statistic (Z-statistic), the effect size and the clusters labelling provided by Johns Hopkins University ICBM-DTI-81 WM labels atlas or MRI Atlas of Human WM, listed on A Appendix.

3.1 Effect of demographics

The goal of this subsection is to determine which demographic variables have a significant effect or trend to statistical significance over FA/MD, using a TFCE-corrected p-value < 0.05 and an uncorrected p-value < 0.01. Analyses were performed for age, gender, IQ (z-scores), YE and handedness. The minimum cluster extend was set as 100 voxels, as mentioned in the methodology. For the continuous demographic variables (age, IQ and YE) the R^2 effect size was reported, while for the categorical variables (gender) the Cohen's d was calculated. Both the effect sizes were calculated based on the uncorrected T-statistic value.

3.1.1 Effects on FA

For the analysis of the effects of demographic variables on FA significant effects were found for the variables age, YE and IQ. Age showed a significant negative effect on FA in anterior corona radiata within a cluster of 40205 voxels. Negative trend effects were exhibit in inferior longitudinal fasciculus, posterior thalamic radiation, corpus callosum and middle cerebellar peduncle. Considering both significant and trend effects it is possible to verify that age has large negative effects widespread across the brain WM. YE showed positive significant effects in superior corona radiata and posterior thalamic radiation. Trend effects for YE were found in corticopontine tract, superior longitudinal fasciculus, posterior thalamic radiation, external capsule and corticospinal tract. Furthermore, IQ revealed significant effects in uncinate fasciculus and posterior corona radiata. Besides significant effects, trend effects for IQ

were found in corpus callosum, corticopontine tract, inferior fronto-occipital fasciculus, anterior corona radiata, limb of internal capsule and superior longitudinal fasciculus.

For the remaining demographics, only trend effects were revealed. Gender showed positive trend effects on FA in middle cerebellar peduncle and cingulum. Actually, this analysis studies the effects of being gender male on FA, because the vector included in the analysis was constructed by 1s and 0s, being 1s set for males and 0s for females. For Handedness no effect survived the minimum cluster extent.

This information is summarized in Table 3.1a and Table 3.1b. In Figures 3.1 , 3.2 , 3.3 and 3.4 the effects of the demographic variables on FA are presented along the brain WM.

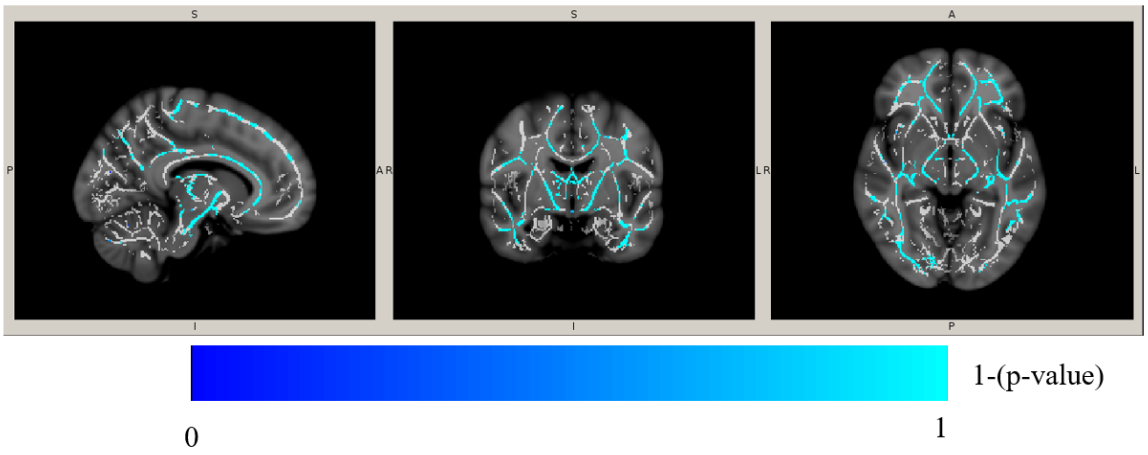

FIGURE 3.1: Visual representation of the WM regions where age has a significant or trend to statistical significance effect on FA. The negative effects of age on FA are presented in blue, the color bar represents the different 1-(p-value) in several shades of blue.

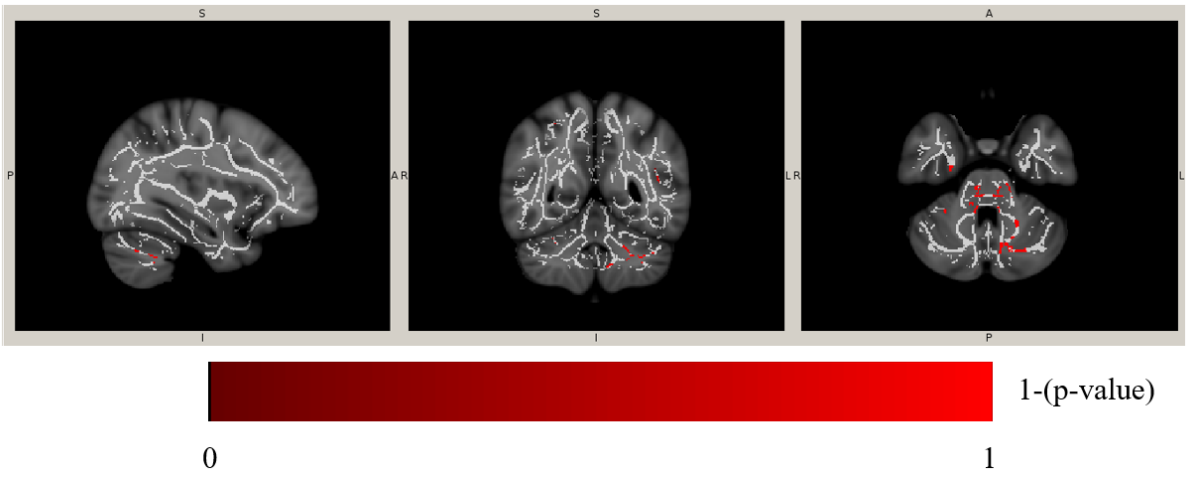

FIGURE 3.2: Visual representation of the regions where gender has a significant or trend to statistical significance effect on FA. The positive effects of gender on FA are presented in red, the color bar represents the different 1-(p-value) in several shades of red.

TABLE 3.1A: TFCE-corrected and uncorrected effects of the demographic variables (Age, Gender, IQ, YE) on FA characterised by the cluster extent (k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). The direction of the effects is represented by (+) or (-) depending if the variable has a positive or negative effect on FA, respectively. TFCE-corrected p-values are marked with **.

Effect of demographic variables on FA							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
Effects of Age (-)							
45843	2.031	1.000×10^{-4}	0.030	131	119	36	Inferior longitudinal fasciculus L
40205	3.636	0.001 **	0.089	113	160	81	Anterior corona radiata L
431	1.113	0.003	0.009	116	45	77	Posterior thalamic radiation L
246	1.919	0.002	0.027	73	38	94	Corpus callosum
200	2.508	0.002	0.045	69	85	45	Middle cerebellar peduncle
170	0.967	0.002	0.007	122	43	83	Inferior longitudinal fasciculus L
150	1.261	0.002	0.012	102	31	80	Posterior thalamic radiation L
149	1.243	0.002	0.011	55	136	117	Superior longitudinal fasciculus R
Effects of gender (+)							
959	0.788	0.003	0.136	123	75	29	Middle cerebellar peduncle
751	1.580	0.003	0.272	99	93	36	Middle cerebellar peduncle
685	2.242	0.002	0.386	77	88	40	Middle cerebellar peduncle
159	1.804	0.005	0.311	69	89	47	Middle cerebellar peduncle
106	3.170	0.003	0.546	65	114	40	Cingulum
Effects of IQ (+)							
2561	1.783	0.001	0.023	125	136	40	Genu of corpus callosum
2554	4.805	0.035 **	0.146	88	153	82	Uncinate fasciculus L
1432	1.573	0.001	0.018	51	149	105	Corticopontine tract R
1022	1.462	0.002	0.016	105	140	62	Inferior fronto-occipital fasciculus L
937	1.616	0.001	0.019	82	97	127	Uncinate fasciculus L
867	1.507	0.001	0.017	132	73	115	Superior longitudinal fasciculus L
595	1.526	0.003	0.017	70	45	105	Corticopontine tract R
413	1.738	0.001	0.022	73	181	60	Inferior fronto-occipital fasciculus L
303	2.477	0.005	0.043	75	138	75	Anterior limb of internal capsule R
272	1.386	0.004	0.014	106	134	121	Inferior fronto-occipital fasciculus L
251	1.363	0.002	0.014	122	175	81	Inferior fronto-occipital fasciculus L
248	1.369	0.002	0.014	118	179	67	Inferior fronto-occipital fasciculus L
173	1.191	0.004	0.010	79	72	126	Corpus callosum
161	1.463	0.002	0.016	102	181	92	Corpus callosum
157	1.731	0.005	0.022	113	110	81	Posterior limb of internal capsule L
141	1.656	0.005	0.020	46	72	98	Posterior thalamic radiation R
137	3.362	0.049 **	0.077	115	80	95	Posterior corona radiata L
125	1.103	0.004	0.009	63	137	101	Superior corona radiata R
119	0.936	0.002	0.006	38	126	63	Inferior longitudinal fasciculus R
115	1.564	0.003	0.018	76	142	61	Inferior fronto-occipital fasciculus R
106	1.216	0.004	0.011	109	78	85	Splenium of corpus callosum
102	1.267	0.003	0.012	121	167	94	Corpus callosum
101	1.826	0.004	0.024	49	101	81	Superior longitudinal fasciculus R

TABLE 3.1B: (Continuation) TFCE-corrected p-values are marked with **. TFCE-corrected and uncorrected effects of the demographic variables (Age, Gender, IQ, YE) on FA characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). The direction of the effects is represented by (+) or (-) depending if the variable has a positive or negative effect on FA, respectively. TFCE-corrected p-values are marked with **.

Effect of demographic variables on FA							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
Effects of YE (+)							
6268	1.620	0.001	0.019	123	93	127	Corticopontine tract L
4136	2.919	0.001	0.059	61	87	102	Corticopontine tract R
3790	3.522	0.024 **	0.084	117	109	102	Superior corona radiata L
2162	2.919	0.025 **	0.059	60	87	89	Posterior thalamic radiation R
814	1.014	0.002	0.008	32	98	60	Superior longitudinal fasciculus R
321	1.859	0.003	0.025	99	148	61	Inferior fronto-occipital fasciculus L
265	4.472	0.045 **	0.128	62	98	118	Posterior thalamic radiation R
182	1.020	0.004	0.008	103	42	104	Posterior thalamic radiation L
174	1.967	0.004	0.028	120	131	80	External capsule L
159	1.517	0.002	0.017	108	55	115	Posterior thalamic radiation L
117	1.467	0.005	0.016	57	131	74	External capsule R
111	1.965	0.004	0.028	134	81	101	Superior longitudinal fasciculus L
101	1.300	0.004	0.012	71	104	138	Corticospinal tract R

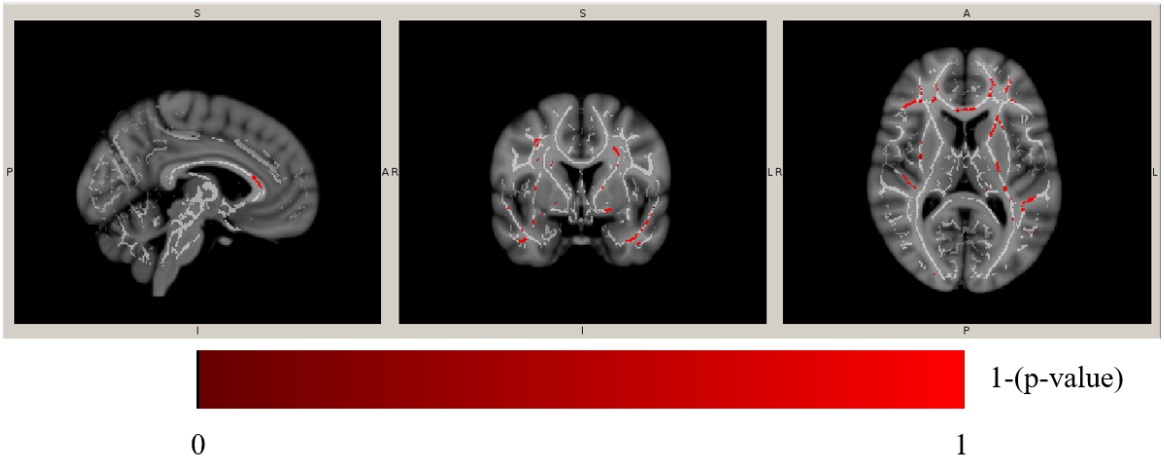

FIGURE 3.3: Visual representation of the regions where IQ has a significant or trend to statistical significance effect on FA. The positive effects of IQ on FA are presented in red, the color bar represents the different 1-(p-value) in several shades of red.

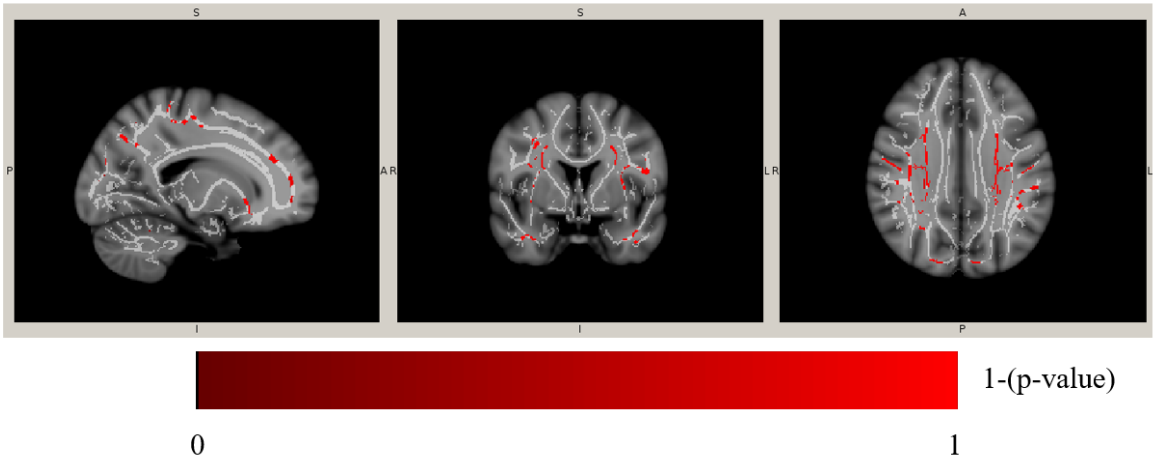

FIGURE 3.4: Visual representation of the regions where YE has a significant or trend to statistical significance effect on FA. The positive effects of YE on FA are presented in red, he color bar represents the different 1-(p-value) in several shades of red.

3.1.2 Effects on MD

For the analysis of the effects of demographic variables on MD only two demographics showed results that survived the minimum cluster extend (Table 3.2). The analyses that resisted were the ones that investigated the effects of gender and IQ on MD. Gender showed significant effects on MD in cerebellar peduncle and trend effects in corpus callosum, cingulum, inferior fronto-occipital fasciculus and superior and inferior longitudinal fasciculus. As previously mentioned the effects of gender are the effects of being male on MD, due to the manner that the vector with the gender information was constructed. There were also significant effects of IQ on MD in anterior thalamic radiation, anterior corona radiata, sagittal stratum and genu of corpus callosum. IQ showed a trend effect on FA in the cingulum. Gender and IQ both revealed to be negatively associated with MD. In Figures 3.5 and 3.6 visual representations of the brain regions where gender and IQ have an effect are presented.

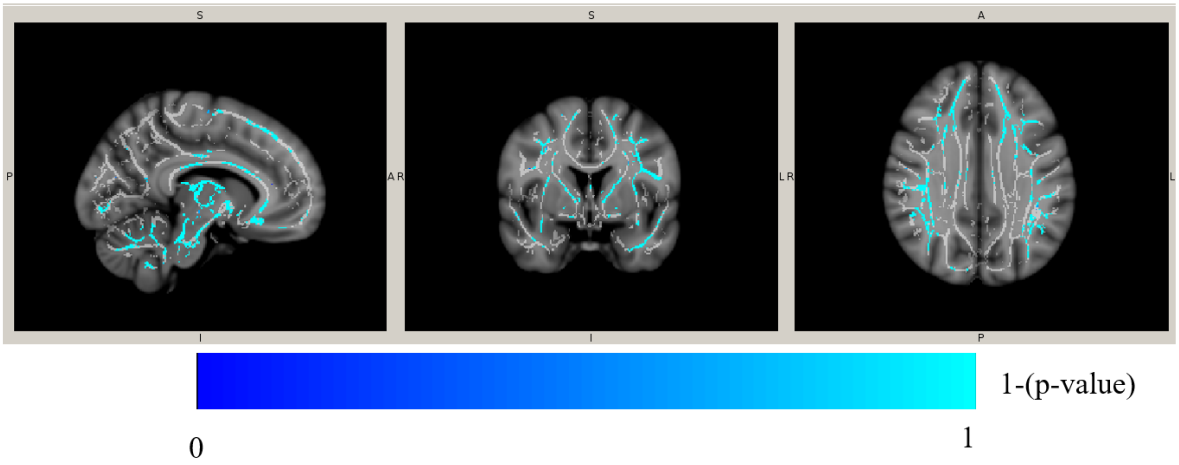

FIGURE 3.5: Visual representation of the WM regions where gender has a significant or trend to statistical significance effect on MD. The negative effects of gender on MD are presented in blue, the color bar represents the different 1-(p-value) in several shades of blue.

TABLE 3.2: TFCE-corrected and uncorrected effects of the demographic variables (gender and IQ) on MD characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). The direction of the effects is represented by (+) or (-) depending if the variable has a positive or negative effect on MD respectively. TFCE-corrected p-values are marked with **.

Effect of demographic variables on MD							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
Effects of gender (-)							
57340	3.945	0.007 **	0.679	99	99	59	Cerebellar peduncle L
44618	1.401	1.000×10^{-4}	0.241	96	76	24	Superior cerebellar peduncle L
5573	1.241	2.000×10^{-4}	0.214	71	43	103	Corpus callosum
889	1.119	0.001	0.193	38	105	103	Superior longitudinal fasciculus R
565	1.652	3.000×10^{-4}	0.284	123	45	89	Inferior fronto-occipital fasciculus L
187	0.783	0.001	0.135	64	104	48	Cingulum
144	1.255	0.003	0.216	140	107	116	Superior longitudinal fasciculus L
125	2.547	0.002	0.438	59	103	50	Cingulum
121	2.354	0.003	0.405	42	109	115	Superior longitudinal fasciculus R
107	1.179	0.001	0.203	67	79	77	Cingulum
Effects of IQ (-)							
1933	3.946	0.004 **	0.102	109	170	72	Anterior thalamic radiation L
1654	4.769	0.029 **	0.142	70	169	74	Anterior corona radiata R
917	4.461	0.041 **	0.127	48	102	60	Sagittal stratum R
176	2.951	0.049 **	0.060	77	156	84	Genu of corpus callosum
103	0.992	0.004	0.007	98	139	100	Cingulum L

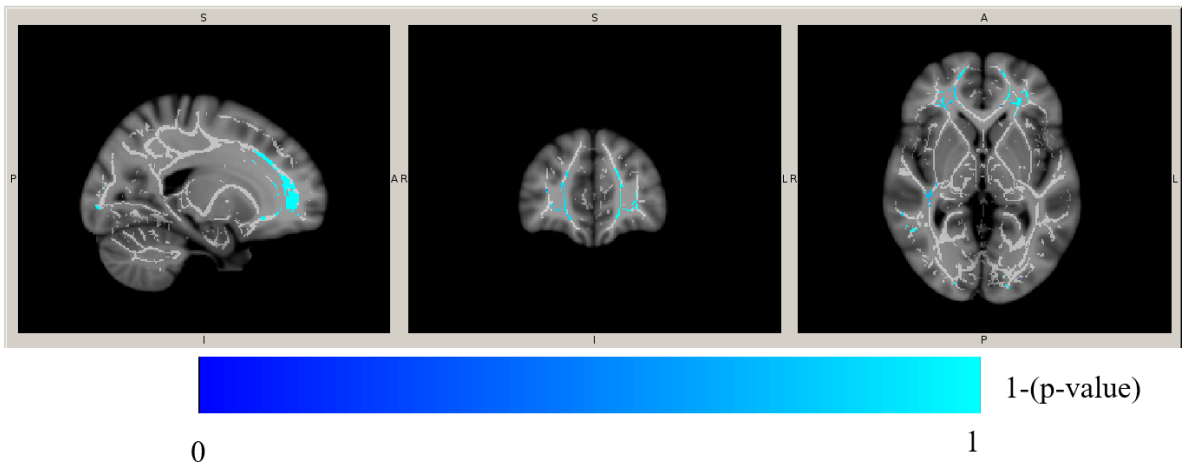

FIGURE 3.6: Visual representation of the regions where IQ has a significant or trend to statistical significance effect on MD. The negative effects of IQ on MD are presented in blue, the color bar represents the different 1-(p-value) in several shades of blue.

3.2 Effect of diagnosis

The purpose of this subsection is to understand if the different groups of diagnosis have an effect on FA/MD, independently of the PRS, using a TFCE-corrected p-value < 0.05 and an uncorrected p-value < 0.01. The minimum cluster extent for this analysis was set to 100 voxels, as referred in the methodology of this work.

In this analysis the main effect of diagnosis on FA/MD was assessed. Afterwards contrasts between the diagnostic groups were also performed to determine the direction of the effects. Age and gender were added to the model because they showed either a significant effect on the diffusion measure (FA/MD) or/and there were significant differences in terms of these variables between diagnostic groups. IQ was not added to the model, because of its intimate relation with the diagnosis. Consequently, YE was also not added to the experimental model because it was not statistically independent from IQ. After inspection using IBM SPSS Statistics 24, YE revealed to be significantly correlated with IQ (r = 0.405, $p = 5.24 \times 10^{-7}$).

For this analysis η^2 values were computed for the main effect of diagnosis based on the uncorrected F-statistic, while cohen's d effect size was calculated for every effect in the post-hoc analysis based on the uncorrected T-statistic.

3.2.1 Effects on FA

The main effect of diagnosis on FA only showed trend effects below the minimum cluster extent in the superior longitudinal fasciculus. Furthermore, the post-hoc analysis revealed a significant effect sustaining that REL are associated with larger values of FA when compared with HC in genu of corpus

callosum. Trend effects for the same contrast were found in middle cerebellar peduncle, corpus callosum, anterior thalamic radiation and inferior fronto-occipital fasciculus and cingulum. Results also demonstrated that REL assume greater FA values when compared with BD in corpus callosum, inferior longitudinal fasciculus, middle cerebellar peduncle, cingulum and inferior fronto-occipital fasciculus, at the uncorrected level. Moreover, trends revealed larger FA values for REL than for SZ in the cingulum and superior longitudinal fasciculus. In this analysis, several effects sustained that REL group is associated with larger FA values than the other three diagnostic groups in several WM regions.This information is summarized in Table 3.3.

The graphical representation presented in Figure 3.8 describes the FA distribuition for the four diagnostic groups in the peak MNI coordinates (80, 147, 90). In Figure 3.7 the WM regions that have significant differences in dignosis on FA are presented.

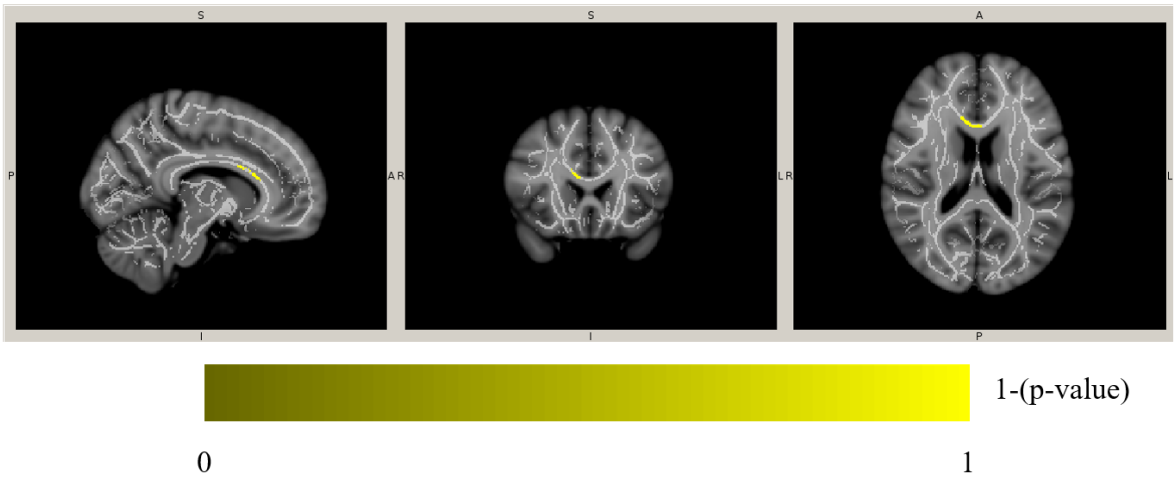

FIGURE 3.7: Visual representation of the regions with significant diferences between diagnosis on FA. The region represented in yellow is located in Genu of Corpus Callosum where REL show larger FA values than HC. The color bar represents the different 1-(p-value) in several shades of yellow.

TABLE 3.3: (TFCE-uncorrected effects of diagnosis on FA for the whole-sample characterised by the cluster extent (k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). TFCE-corrected p-values are marked with **.

Effect of diagnosis on FA							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
REL >HC							
466	2.508	0.002	0.430	129	70	39	Middle cerebellar peduncle
462	0.674	0.001	0.116	63	83	44	Middle cerebellar peduncle
406	1.574	0.002	0.270	49	141	97	Middle cerebelar peduncle
393	1.078	0.003	0.185	102	122	107	Body of corpus callosum
323	1.085	0.002	0.186	106	36	93	Corpus callosum
201	4.588	0.044**	0.787	80	147	90	Genu of corpus callosum
199	0.973	0.003	0.167	113	80	71	Cingulum L
193	1.338	0.004	0.229	79	42	73	Corpus callosum
150	1.062	0.003	0.182	83	104	74	Anterior thalamic radiation R
150	1.110	0.003	0.190	54	68	28	Middle cerebelar peduncle
128	1.711	0.004	0.293	54	172	66	Inferior fronto-occipital fasciculus R
106	1.428	0.004	0.245	49	152	66	Inferior fronto-occipital fasciculus R
102	1.049	0.004	0.180	79	73	128	Corpus callosum
REL >BD							
305	1.263	0.002	0.217	46	70	65	Inferior longitudinal faciculus R
225	1.214	0.003	0.208	102	122	107	Body of corpus callosum
196	1.174	0.001	0.201	69	89	47	Middle cerebellar peduncle
191	0.948	0.001	0.163	58	103	43	Cingulum R
186	2.897	0.003	0.497	45	115	36	Inferior longitudinal fasciculus R
140	2.095	0.003	0.359	39	129	60	Inferior longitudinal faciculus R
139	0.943	0.004	0.162	73	41	67	Inferior fronto-occipital fasciculus R
104	1.482	0.006	0.254	73	109	129	Corpus callosum
REL >SZ							
353	0.943	0.001	0.162	115	115	39	Cingulum L
257	1.331	0.003	0.228	35	76	75	Superior longitudinal fasciculus R
140	1.655	0.002	0.284	58	77	125	Superior longitudinal fasciculus R

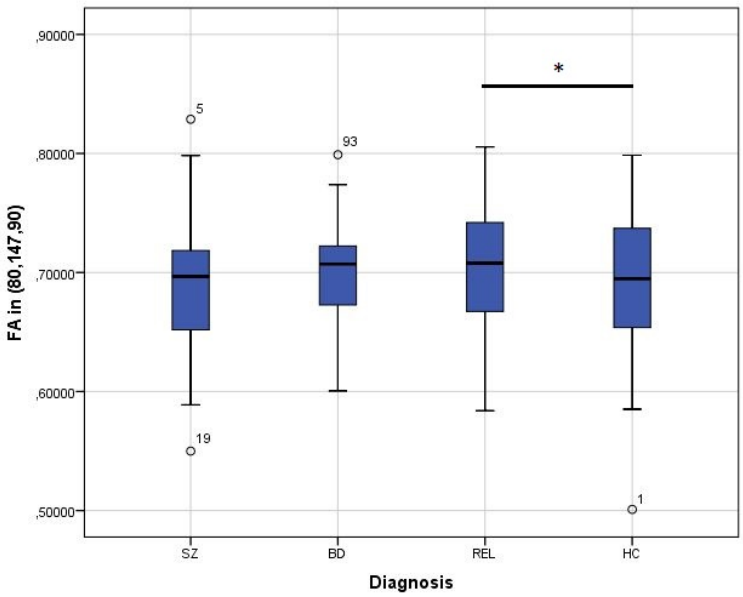

FIGURE 3.8: Box plot of the FA distribuition of the diagnostic groups [schizophrenia(SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (80, 147, 90) for the whole-sample. The groups with significant differences are poited with *.

3.2.1.1 Participants subsample

Based on the previous sections, is possible to verify that age has a significant effect on FA and there are significant differences between diagnostic groups in terms of age. This subsection has the goal of finding if the diagnosis influences FA, while controlling for the confounder age.

The analysis of the main effect of diagnosis on FA for the age-matched sample (n = 115) showed largest effects when compared with the effects obtained for the whole-sample (n = 143). However, the effects obtained in the age-matched sample didn't reach significance. The largest trend showed for the analysis of the main effect of diagnosis was in the inferior longitudinal fasciculus within a cluster of 880 voxels. Moreover, there were also trend effects in superior and inferior longitudinal fasciculus. corpus callosum, external capsule, cingulum and corticospinal tract. In Figure 3.9, the main effects of diagnosis on FA are represented along the brain WM.

Furthermore, the post-hoc analysis revealed that REL are associated with larger values of FA when compared with BD in posterior corona radiata, at the significance level. For this contrast, trend effects were found in inferior cerebellar peduncle, cingulum and corticopontine tract. The analysis also showed trend effects sustaining that REL have greater FA values when compared with SZ group in corticospinal tract, superior and inferior longitudinal fasciculus, cingulum, inferior fronto-occipital fasciculus, corpus callosum and uncinate fasciculus. At last, the results showed that REL have larger FA values than HC reaching significance in posterior corona radiata and corticospinal tract. Trends for this contrast were found in superior and inferior longitudinal fasciculus, corticospinal tract and cingulum.

The information about the effects of diagnosis on FA in the age-matched sample is summarized in Table 3.4a and Table 3.4b. A representation of the FA distribuition in the four diagnostic groups in all the significant effects peak MNI coordinates is presented in Figures 3.10, 3.11 and 3.12.

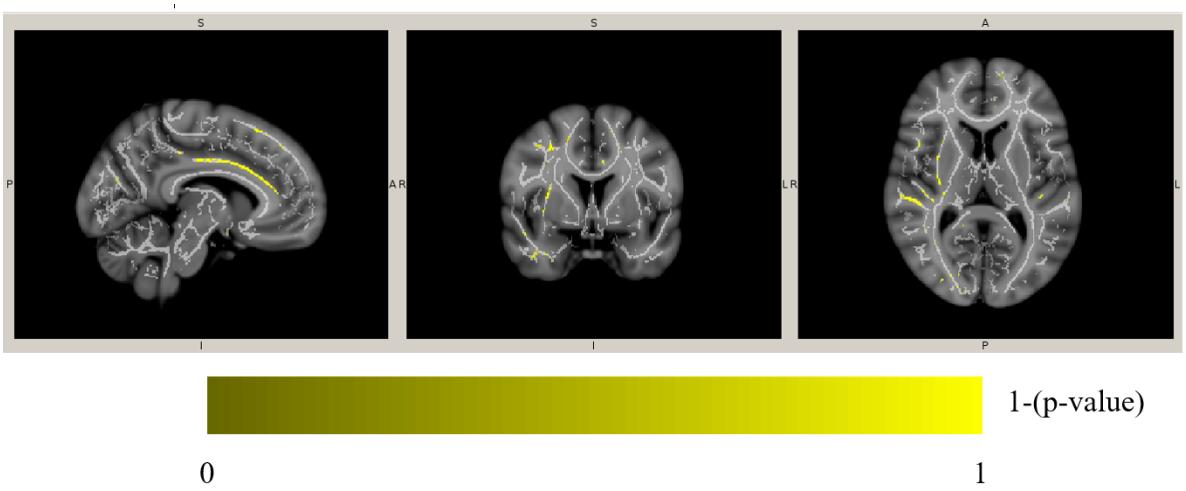

FIGURE 3.9: Main effect of diagnosis on FA for a TFCE-uncorrected p-value < 0.01, across the participants age-matched subsample. In yellow are represented the regions with trend effects.The color bar represents the different 1-(p-value) in several shades of yellow.

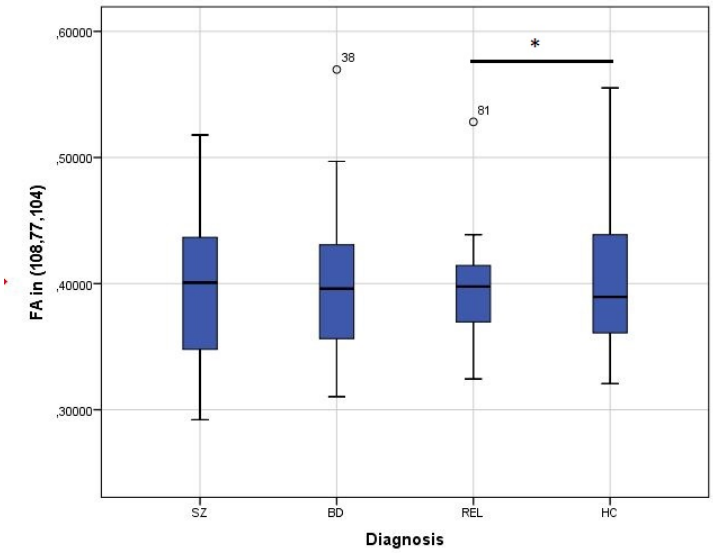

FIGURE 3.10: Box plot of the FA distribuition of the diagnostic groups [szchizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (108, 77, 104) for the participants subsample. The groups with significant differences are poited with *.

TABLE 3.4A: TFCE-uncorrected effects of diagnosis on FA for the participants subsample characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). TFCE-corrected p-values are marked with **.

Effect of diagnosis on FA							
k	Z-statistic	p-value	Effect size	Peak MNI coordinates			Cluster labelling
				x{mm}	y{mm}	z{mm}	
Main effect of diagnosis							
880	1.917	0.001	0.049	41	128	56	Inferior longitudinal faciculus R
831	1.666	0.001	0.043	67	75	70	Cingulum R
728	2.025	0.001	0.052	81	50	94	Cingulum R
491	3.202	4.000×10^{-4}	0.080	98	162	85	Cingulum L
360	3.413	9.997×10^{-4}	0.084	38	78	67	Superior longitudinal fasciculus R
312	3.428	0.003	0.085	139	76	102	Superior longitudinal fasciculus L
287	2.103	2.000×10^{-4}	0.054	61	39	77	Inferior longitudinal faciculus R
221	1.787	0.001	0.046	99	145	129	Inferior fronto-occipital fasciculus L
205	1.608	0.004	0.042	61	132	88	External capsule R
195	2.044	0.002	0.052	43	139	76	Superior longitudinal fasciculus R
166	2.465	0.004	0.062	65	74	82	Splenium of corpus callosum
139	4.490	0.003	0.108	49	91	105	Superior longitudinal fasciculus L
133	2.039	0.002	0.052	58	136	120	Inferior fronto-occipital fasciculus R
127	1.931	0.003	0.050	73	105	132	Corpus callosum
121	2.278	0.003	0.058	108	78	127	Corpus callosum
111	1.948	0.003	0.050	122	97	117	Corticospinal tract L
107	1.525	0.004	0.040	141	92	112	Superior longitudinal fasciculus L
REL >BD							
56273	1.266	1.000×10^{-4}	0.244	73	54	35	Inferior cerebellar peduncle R
47270	2.892	0.004 **	0.556	108	77	104	Posterior corona radiata L
445	1.294	3.000×10^{-4}	0.249	65	114	40	Cingulum R
195	0.650	0.002	0.125	79	144	129	Corticopontine tract R
REL >SZ							
760	1.175	0.002	0.226	98	93	35	Corticospinal tract L
590	1.302	0.002	0.251	59	130	116	Superior longitudinal fasciculus R
553	1.026	0.003	0.197	76	54	88	Cingulum R
398	1.610	0.003	0.310	39	110	115	Superior longitudinal fasciculus R
340	1.914	0.003	0.368	73	41	67	Inferior fronto-occipital fasciculus R
261	1.342	0.004	0.258	102	44	67	Inferior fronto-occipital fasciculus L
243	1.004	0.003	0.193	98	163	83	Cingulum L
239	1.392	0.003	0.268	96	111	108	Cingulum L
232	1.139	0.003	0.219	54	93	49	Inferior longitudinal faciculus R
225	1.387	0.004	0.267	52	138	43	Uncinate fasciculus R
201	1.177	0.004	0.227	83	57	106	Corpus callosum
179	1.662	0.003	0.320	55	56	111	Superior Longitudinal Fasciculus R
159	0.898	0.003	0.173	61	87	104	Superior Longitudinal Fasciculus R
154	1.199	0.004	0.231	128	141	109	Superior Longitudinal Fasciculus L
151	0.940	0.005	0.181	101	71	47	Inferior Cerebellar Peduncle L
150	1.246	0.003	0.240	141	84	59	Superior Longitudinal Fasciculus L
141	1.448	0.005	0.279	117	124	35	Cingulum L
118	1.324	0.005	0.255	47	80	116	Superior Longitudinal Fasciculus R
111	1.797	0.005	0.346	47	104	81	Superior Longitudinal Fasciculus R
111	1.556	3.000×10^{-4}	0.299	62	39	77	Inferior Longitudinal Faciculus R
105	1.615	0.005	0.311	133	136	50	Uncinate fasciculus L

TABLE 3.4B: (Continuation) TFCE-uncorrected effects of diagnosis on FA for the participants subsample characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). TFCE-corrected p-values are marked with **.

Effect of diagnosis on FA							
k	Z-statistic	p-value	Effect size	Peak MNI coordinates			Cluster labelling
				x{mm}	y{mm}	z{mm}	
HC >BD							
5155	1.020	3.000×10^{-4}	0.196	37	78	63	Superior longitudinal fasciculus R
4742	3.893	0.039 **	0.749	61	70	95	Posterior corona radiata R
828	2.047	0.001	0.394	73	109	127	Corticospinal tract R
701	3.580	0.043 **	0.689	67	92	122	Corticospinal tract R
601	1.307	0.002	0.252	109	92	139	Corticospinal tract L
490	2.370	0.002	0.456	36	122	69	Inferior longitudinal faciculus R
400	2.227	0.002	0.429	42	129	104	Superior longitudinal fasciculus R
332	1.209	0.002	0.233	101	112	139	Corticospinal tract L
147	1.396	0.003	0.269	148	97	62	Superior longitudinal fasciculus L
133	1.219	0.005	0.235	150	102	80	Inferior longitudinal faciculus L
116	2.100	0.005	0.404	79	103	105	Cingulum R
109	2.163	0.003	0.416	38	105	98	Superior longitudinal fasciculus R

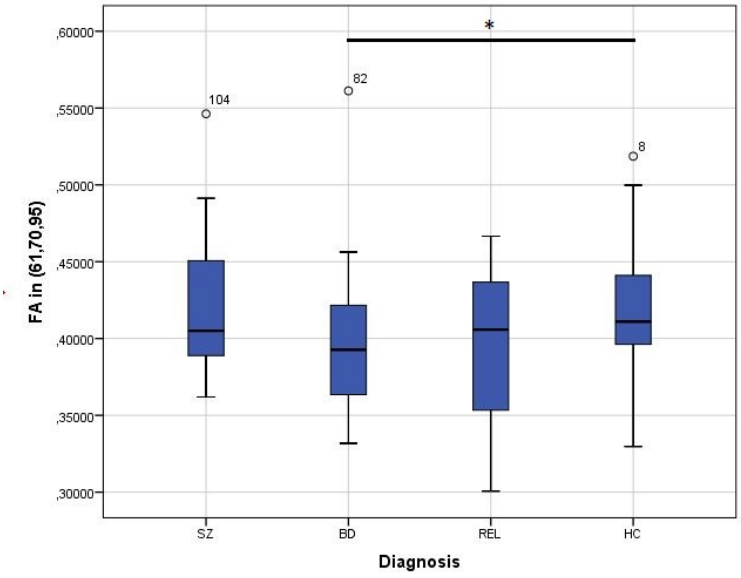

FIGURE 3.11: Box plot of the FA distribuition of the diagnostic groups [schizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (61, 70, 95) for the participants subsample. The groups with significant differences are poited with *.

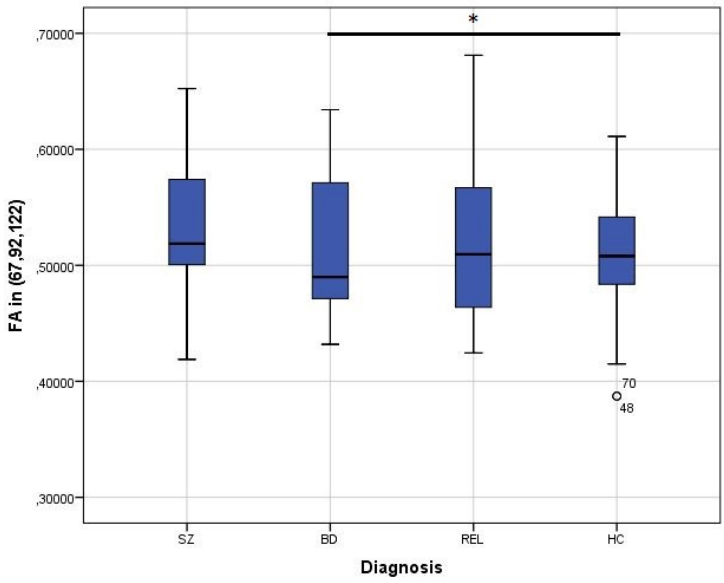

FIGURE 3.12: Box plot of the FA distribuition of the diagnostic groups [schizophrenia(SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (67, 92, 122) for the participants subsample. The groups with significant differences are poited with *.

3.2.2 Effects on MD

For the analysis of the main effects of diagnosis on MD no effect survived the minimum cluster extend of 100 voxels. The largest effect reported had a cluster extent of 10 voxels located in middle cerebellar peduncle. Furthermore, post-hoc analysis showed effects in which BD group assumes larger MD values than HC and REL in several WM regions. Trend effects were found in anterior thalamic radiation and inferior fronto-occipital fasciculus, regions in which BD is associated with greater values of MD when compared with HC. Moreover, significant effects showed that the effect of BD is associated with larger MD values than REL in superior longitudinal fasciculus and anterior thalamic radiation within clusters with 20345 and 273 voxels, respectively. Trend effects were also found in inferior and superior longitudinal fasciculus, corpus callosum, inferior fronto-occipital fasciculus, posterior thalamic radiation. external and internal capsule and uncinate fasciculus. These results are summarized in Table 3.5 . In Figures 3.14 and 3.15 are represented the MD distribuitions in the different diagostic groups in the regions with significant effect. In Figure 3.13 the significant effects of diagnosis on MD, for the participant subsample, are presented in yellow.

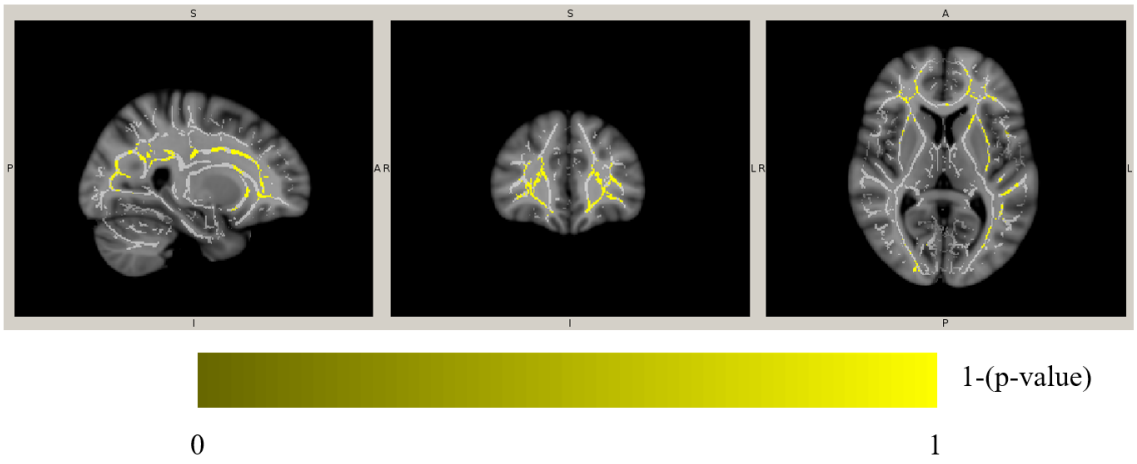

FIGURE 3.13: Visual representation of the regions with significant diferences between diagnosis on MD. In the region represented in yellow BD patients have larger values of MD when compared with REL.The color bar represents the different 1-(p-value) in several shades of yellow.

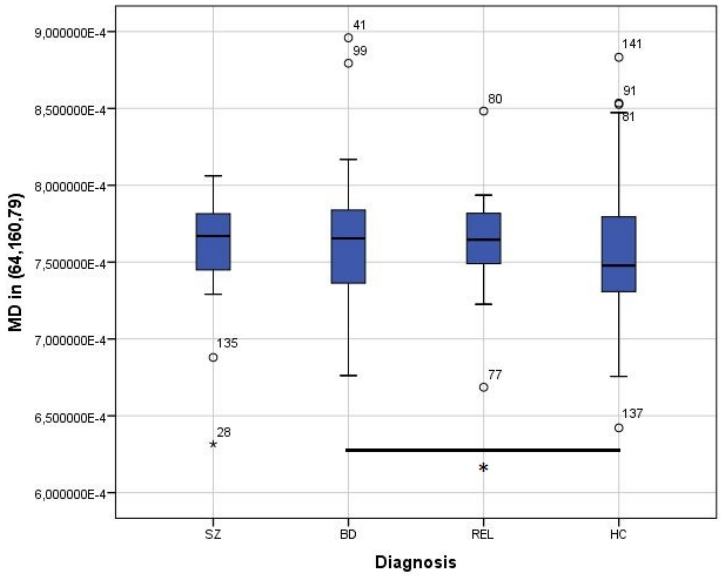

FIGURE 3.14: Box plot of the MD distribuition of the diagnostic groups [schizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (64, 160, 79). The groups with significant differences are poited with *.

TABLE 3.5: TFCE-corrected and uncorrected effects of diagnosis on MD for the whole-sample characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left). TFCE-corrected p-values are marked with **.

Effect of diagnosis on MD							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
BD >HC							
218	1.300	0.003	0.223	81	118	78	Anterior thalamic radiation R
152	1.527	0.001	0.262	82	138	58	Anterior thalamic radiation R
111	1.455	0.001	0.250	115	129	61	Inferior fronto-occipital fasciculus L
BD >REL							
20345	3.356	0.038 **	0.576	64	160	79	Superior longitudinal fasciculus R
1307	1.918	0.002	0.329	110	77	82	Splenium of corpus callosum
1020	3.194	0.002	0.548	72	38	75	Corpus callosum
731	1.678	0.002	0.288	137	111	37	Inferior longitudinal faciculus L
456	2.237	0.004	0.384	81	119	101	Body of corpus callosum
347	1.171	0.001	0.201	105	35	72	Inferior fronto-occipital fasciculus L
273	0.499	0.049 **	0.086	67	138	87	Anterior thalamic radiation R
257	3.021	0.003	0.518	120	107	82	External capsule L
232	1.917	0.001	0.329	115	129	61	Inferior fronto-occipital fasciculus L
227	1.303	0.004	0.223	116	133	60	Uncinate fasciculus L
208	1.548	0.001	0.265	110	89	79	Posterior thalamic radiation L
204	1.403	0.003	0.241	99	185	65	Corpus callosum
189	1.923	0.004	0.330	99	70	89	Cingulum L
185	1.978	0.004	0.339	146	120	87	Superior longitudinal fasciculus L
184	1.677	0.001	0.288	125	45	84	Superior longitudinal fasciculus L
150	2.013	0.003	0.345	127	152	63	Uncinate fasciculus L
132	1.373	0.002	0.236	50	150	62	Inferior fronto-occipital fasciculus R
126	1.727	0.003	0.296	36	88	100	Superior longitudinal fasciculus R
116	1.773	0.006	0.304	59	129	84	External capsule R
112	1.730	0.004	0.297	106	138	75	Anterior limb of internal capsule L
112	1.056	0.002	0.181	125	116	41	Inferior longitudinal faciculus L
109	2.607	0.006	0.447	65	70	100	Inferior fronto-occipital fasciculus R
109	1.679	0.005	0.288	140	82	93	Superior longitudinal fasciculus L
109	1.474	0.003	0.253	128	98	50	Inferior longitudinal faciculus L
101	2.514	0.006	0.431	90	145	90	Body of corpus callosum

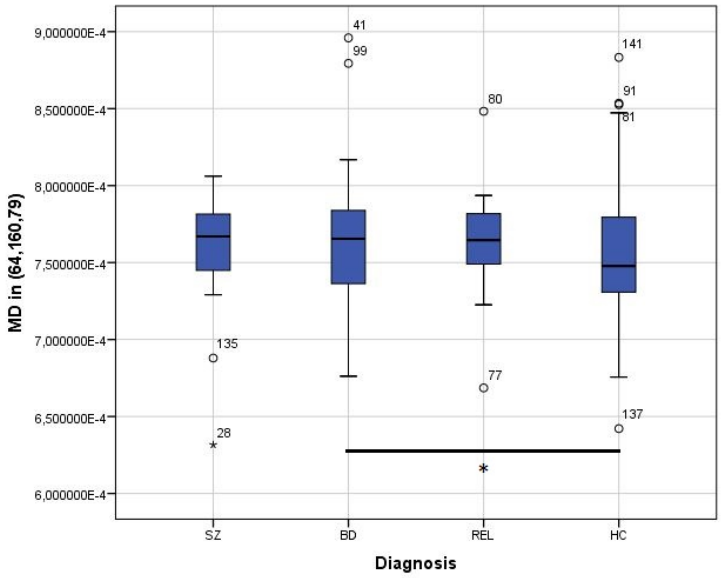

FIGURE 3.15: Box plot of the MD distribuition of the diagnostic groups [szchizophrenia (SZ), bipolar disorder (BD), relatives (REL) of SZ and BD and healthy controls (HC)] in (67, 138, 87). The groups with significant differences are poited with *.

3.3 Effect of PRS

The effects of PRS on FA/MD are reported along this section. The effects presented are independent from the diagnostic group and for a TFCE-corrected p-value < 0.05 and an uncorrected p-value < 0.01. No minimum cluster extent was considered in this analysis.

Age and gender were added to the experimental design because both these variables had a significant effect on FA or MD. Besides that, the great effects of these variables on the diffusion measures have been reported in previous studies. R^2 effect size was computed for each effect based on the uncorrected T-statistic.

3.3.1 Effects on FA

For the analysis of the effects of PRS on FA only small trend effects were found (Table 3.6). The positive effects of PRS were larger than the negative effects and were located in uncinate fasciculus, corpus callosum, middle cerebellar peduncle and anterior thalamic radiation. Negative trend effects of PRS on FA were located in superior longitudinal fasciculus, posterior and anterior thalamic radiation, corticopontine tract, middle cerebellar peduncle and cingulum. This information is summarized in Table 3.6. In Figure 3.16, the WM regions of the trend effects are presented.

TABLE 3.6: TFCE-uncorrected positive and negative effects of PRS on FA characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left).

Effect of PRS on FA							
k	**Z-statistic**	**p-value**	**Effect size**	**Peak MNI coordinates**			**Cluster labelling**
				x{mm}	**y{mm}**	**z{mm}**	
Positive effects							
42	2.618	0.001	0.048	122	142	69	Uncinate fasciculus L
2	3.395	0.007	0.079	111	177	92	Corpus callosum
1	3.356	9.999×10^{-4}	0.077	111	79	46	Middle cerebellar peduncle
1	2.665	0.005	0.050	103	135	64	Anterior thalamic radiation L
1	1.411	0.010	0.010	125	141	64	Uncinate fasciculus L
1	2.517	0.008	0.008	69	183	79	Middle cerebellar peduncle
Negative effects							
17	1.193	0.008	0.010	57	84	121	Superior longitudinal fasciculus R
13	1.112	0.006	0.009	58	85	112	Superior longitudinal fasciculus R
8	2.984	0.001	0.062	98	97	67	Posterior thalamic radiation L
3	2.623	0.002	0.049	94	100	67	Anterior thalamic radiation L
2	3.651	4.000×10^{-4}	0.090	108	129	75	Corticopontine tract L
1	1.588	0.007	0.018	72	75	47	Middle cerebellar peduncle
1	1.301	0.009	0.012	73	74	49	Middle cerebellar peduncle
1	3.011	0.008	0.063	89	99	68	Superior cerebellar peduncle
1	2.749	0.005	0.053	144	75	72	Superior longitudinal fasciculus L
1	2.746	0.004	0.053	75	80	80	Cingulum
1	2.589	0.010	0.047	48	103	101	Superior longitudinal fasciculus R
1	2.775	0.003	0.054	61	142	101	Anterior corona radiata R

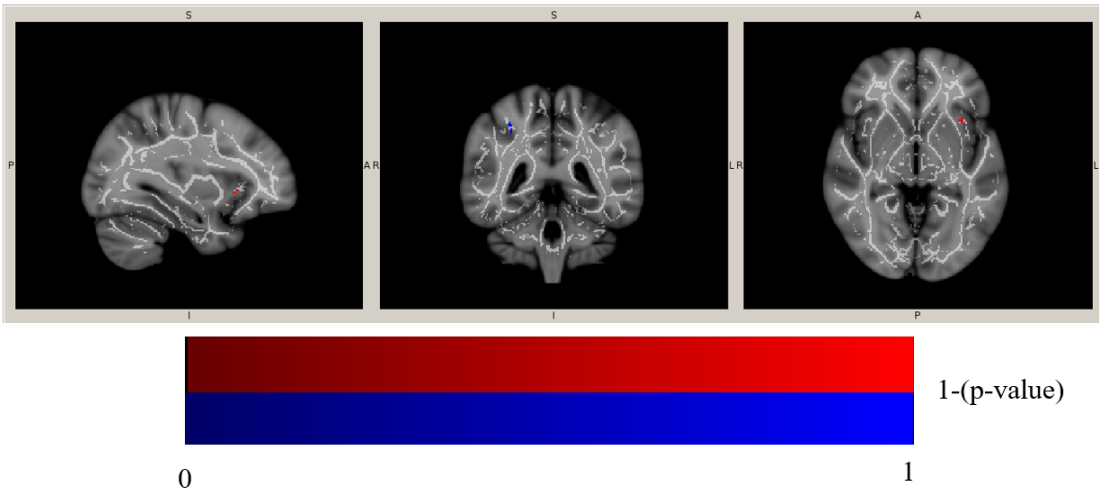

FIGURE 3.16: Visual representation of the effects of PRS on FA for a TFCE-uncorrected p-value < 0.01, across all participants. The regions that correspond to positive effects are presented in red and the regions of the negative effects are presented in blue. The color bars represent the different 1-(p-value) in several shades of blue and red.

3.3.2 Effects on MD

For the analysis of the effects of PRS on MD only trend effects were found. The positive effects were slightly smaller than the negative effects and were located in inferior longitudinal fasciculus and posterior corona radiata. Negative effects of PRS on MD were all located in inferior cerebellar peduncle. This information is summarized in Table 3.7. In Figure 3.17 a visual representation of the brain the negative and positive effects of PRS on FA are presented.

TABLE 3.7: TFCE-uncorrected positive and negative effects of PRS on MD characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left).

Effect of PRS on MD							
k	Z-statistic	p-value	Effect size	Peak MNI coordinates			Cluster labelling
				x{mm}	y{mm}	z{mm}	
Positive effects							
3	2.736	0.005	0.053	122	81	68	Inferior longitudinal fasciculus L
1	3.108	0.001	0.067	67	84	105	Posterior corona radiata R
Negative effects							
13	2.970	0.001	0.061	90	73	48	Inferior cerebellar peduncle
1	2.612	0.008	0.048	88	70	51	Inferior cerebellar peduncle

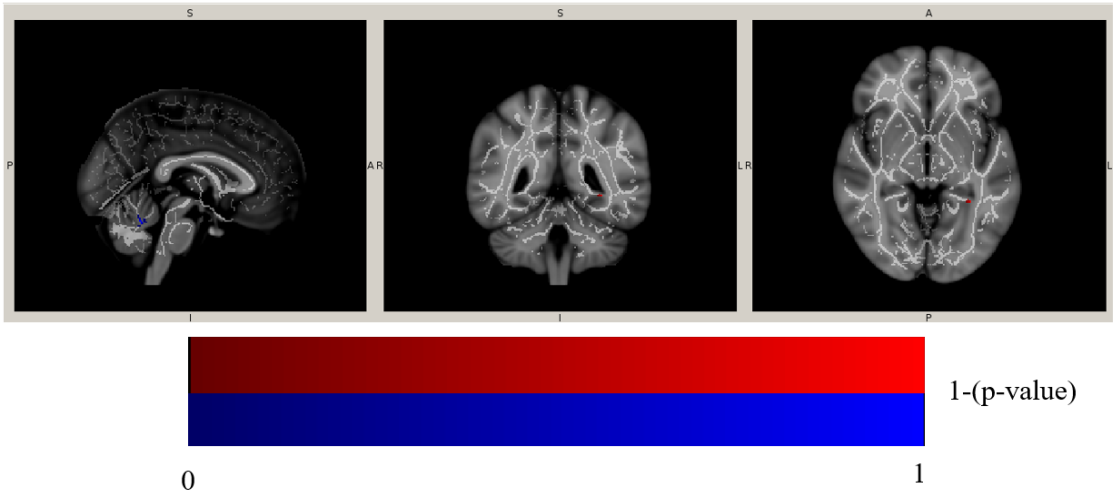

FIGURE 3.17: Visual representation of the effects of PRS on MD for a TFCE-uncorrected p-value < 0.01, across all participants. The regions that correspond to positive effects are presented in red and the regions of the negative effects are presented in blue. The color bars represent the different 1-(p-value) in several shades of blue and red.

3.4 Interaction effect PRS x Diagnosis

In this section the results of the interaction analysis between PRS and diagnosis are presented for a TFCE-corrected p-value < 0.05 and an uncorrected p-value < 0.01. For practical purposes only clusters

with more than 50 voxels were reported in the results section. Age and gender were added to the analysis model because both these variables have a significant effect in at least one of the diffusion measures. R^2 effect size was computed for each effect based on the uncorrected T-statistic.

The effects with less than 50 voxels are reported in B Appendix, characterised by their cluster size, 1-(p-value) and peak coordinates.

3.4.1 Effects on FA

The interaction effects between PRS x diagnosis are summarized in Table 3.8. Trend effects showed that the correlation between PRS and FA is greater in BD when compared with HC in middle cerebellar peduncle, superior and inferior longitudinal fasciculus, cingulum, corpus callosum, inferior fronto-occipital fasciculus, cortical spinal tract and uncinate fasciculus. Trends also exhibit larger correlation of PRS with FA for REL than HC in inferior fronto-occipital fasciculus. At last, trends showed that the correlation of PRS and FA was smaller for SZ than for REL and BD in middle cerebellar peduncle. In Figure 3.18 are presented the regions of the PRS x diagnosis interaction effects that had more than 50 voxels.

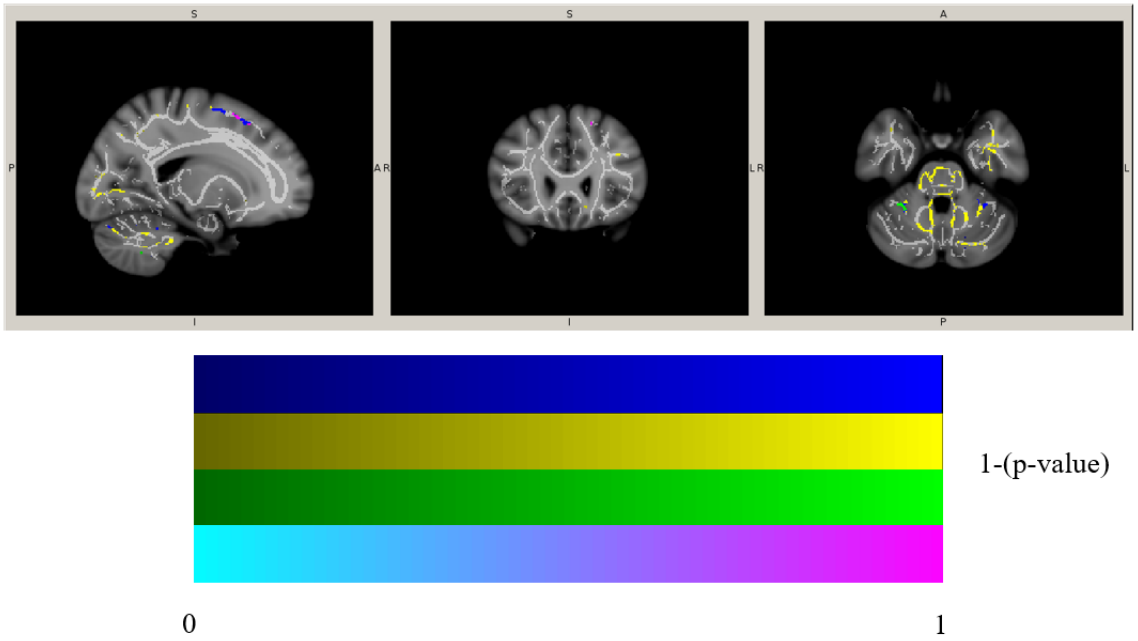

FIGURE 3.18: Visual representation of the PRS x diagnosis interaction effects on FA for a TFCE-uncorrected p-value < 0.01, across all participants. The regions that correspond to (1) the contrast PRS in BD > PRS in HC are presented in yellow; (2) the contrast PRS in REL > PRS in HC are presented in pink; (3) the contrast PRS in BD > PRS in SZ are presented in blue; (4) the contrast PRS in REL > PRS in SZ is presented in green. The color bars represent the different 1-(p-value) in several shades of blue, pink, green and yellow.

TABLE 3.8: TFCE- uncorrected PRS x diagnosis interaction effects on FA characterised by the cluster extent(k), z-statistic, p-value, effect size, MNI coordinates and WM label(R-right; L-left).

PRS x diagnostic effects on FA							
k	Z-statistic	p-value	Effect size	Peak MNI coordinates			Cluster labelling
				x{mm}	y{mm}	z{mm}	
PRS in BD >PRS in HC							
2828	1.379	0.001	0.014	77	76	35	Middle cerebellar peduncle
362	2.312	0.003	0.039	147	85	85	Superior longitudinal fasciculus L
299	1.608	0.002	0.019	122	104	43	Cingulum
213	1.230	0.006	0.011	106	48	46	Middle cerebellar peduncle
200	1.404	0.004	0.015	87	59	46	Inferior cerebellar peduncle
167	1.074	0.002	0.009	116	73	48	Middle cerebellar peduncle
157	1.462	0.001	0.016	67	83	45	Middle cerebellar peduncle
119	2.334	0.006	0.040	93	99	48	Middle cerebellar peduncle
117	1.113	0.004	0.009	105	31	75	Inferior fronto-occipital fasciculus L
104	1.205	0.004	0.011	103	47	72	Inferior fronto-occipital fasciculus L
93	1.139	0.004	0.010	99	92	134	Corpus callosum
89	0.790	0.002	0.005	99	115	136	Corpus callosum
76	1.453	0.004	0.016	101	100	141	Cortical spinal tract L
74	0.943	0.004	0.007	107	54	115	Inferior longitudinal fasciculus L
71	2.447	0.005	0.043	97	147	58	Uncinate fasciculus L
68	2.084	0.004	0.032	81	140	131	Corpus callosum
65	1.571	0.005	0.018	130	145	106	Inferior fronto-occipital fasciculus L
64	0.676	0.001	0.003	106	130	130	Corpus callosum
57	1.655	0.005	0.020	113	63	123	Superior longitudinal fasciculus L
PRS in REL >PRS in HC							
81	0.949	0.001	0.007	108	150	124	Inferior fronto-occipital fasciculus R
PRS in BD >PRS in SZ							
153	1.069	0.002	0.009	60	85	39	Middle cerebellar peduncle
76	1.526	0.004	0.017	112	88	45	Middle cerebellar peduncle
67	1.603	0.002	0.019	107	131	131	Corticopontine tract L
54	0.217	0.004	0.000	108	150	124	Corticopontine tract L
53	1.511	0.004	0.017	69	85	45	Middle cerebellar peduncle
PRS in REL >PRS in SZ							
122	1.440	0.003	0.015	57	81	42	Middle cerebellar peduncle

3.4.1.1 Participants subsample

Regarding the interaction analysis between PRS and FA for the age-matched sample no effect was larger than 50 voxels. The effect that was closer to reaching this value showed that the correlation of PRS with FA was larger in SZ than BD in right superior longitudinal fasciculus within a cluster of 38 voxels.

3.4.2 Effects on MD

The interaction effects between PRS and diagnosis on MD showed that the correlation between PRS and MD was larger in HC than in BD in inferior fronto-occipital fasciculus (108, 145, 125) within a cluster with 70 voxels (t = 1.199; p-value = 0.003). The effect size revealed that this effect has medium

strength (R^2 = 0.011). In Figure 3.19 the PRS x diagnosis interaction effects on MD are presented along the brain WM.

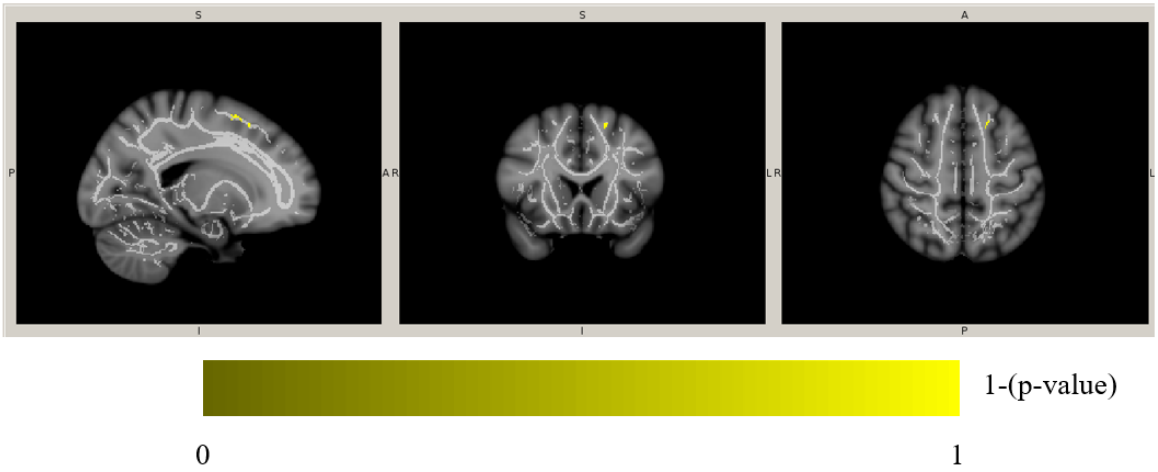

FIGURE 3.19: Visual representation of the PRS x diagnosis interaction effects on MD for a TFCE-uncorrected p-value < 0.01, across all participants. The regions in yellow correspond to the contrast PRS in HC > PRS in BD. The color bars represent the different 1-(p-value) in several shades of yellow.

4 Discussion

The purpose of this work was to assess the effects of a SZ PRS, which compounds recently-associated SZ risk variants, on brain WM microstructure, in a Caucasian sample divided in four diagnostic groups (SZ, BD, REL and HC). For completeness, the effect of several demographic variables and the effect of the diagnostic groups on WM microstructure were also studied. Based on above-mentioned literature, my hypothesis is that a higher PRS will have an influence on WM microstructure. Moreover, that the effect of the PRS is negatively associated with WM microstructure, being this effect different in the several diagnositc groups.

In this study, several significant effects on FA were found (TFCE-corrected p< 0.05): (1) an effect of age, IQ and YE, (2) a significant difference between REL and HC, (3) a significant difference between REL and BD and (4) a significant difference between HC and BD. However, no significant effects of SZ PRS on FA were found. Moreover, no significant effects of PRS considering four diagnostic groups on FA were found. Furthermore, some significant effects on MD were found (TFCE-corrected p < 0.05): (1) an effect of gender and IQ and (2) a significant difference between BD and REL. As for PRS, no significant effects on MD were found, whether considering or not the four diagnostic groups.

In the following subsections, the effects obtained in this study are discussed in more detail.

4.1 Effect of demographics

Several demographic variables revealed to significantly influence FA/MD. The effects of the demographic variables on FA/MD are important to the reliability of the following analyses. The assessment of these effects is important, so it is possible to subtract the contribution of demographic variables that have large effects on FA/MD, while studying the effects of other variables.

The results showed that age had significant negative influence on FA in anterior corona radiata with a medium effect size (R^2=0.089). Trend effects were also evidenced in other WM regions as middle cerebellar peduncle, corpus callosum and superior and inferior longitudinal fasciculus. These results are supported by the findings in previous studies Across studies, the predominant findings are a decrease in FA in function of increasing adult age, suggesting an age-related decline in the integrity of WM microstructure [90], [91], [93]. Furthermore, previous literature has stated that frontal WM is more vulnerable to damage than the remaining brain [93]. The large significant effect of age in anterior corona radiata (frontal WM region) upholds this idea. DTI studies have included SZ and BD patients with ages

from about 16 to 60 years [66], [69], [74] which is similar to the age range of this study 18 to 63 years. This age range permits the study of the influence of age on brain white matter microstructure. Gender revealed to have trend tostatistical significant positive effects on FA in middle cerebellar peduncle and cingulum. Clarifying, being male has a positive influence on FA in the regions previously enumerated. Furthermore, the results of this work show that males have negative significant influence on MD in cerebellar peduncle. Trends were also found in regions as the cingulum and cerebellar peduncle. Prior literature has also revealed that males have higher FA than females in the midcingulum bundle while controlling for differences in brain volume [116]. The influence of males on MD was expected, as FA and MD have antagonistic behaviour.

The results showed that IQ is significantly correlated with FA and MD. IQ showed a positive correlation with FA in uncinate fasciculus (R^2=0.146) and posterior corona radiata (R^2=0.077) and a negative correlation with MD in anterior thalamic radiation (R^2=0.102), anterior corona radiata (R^2=0.142), sagittal stratum (R^2=0.127) and genu of corpus callosum (R^2=0.060). For FA, trend effects of IQ have been found in corpus callosum, corticopontine tract, inferior fronto-occipital fasciculus, anterior corona radiata, limb of internal capsule and superior longitudinal fasciculus. A previous study has shown that IQ has a positive correlation with FA, mainly in the corpus callosum, in a sample with children (12 – 15 years) which upholds the trend effects of IQ in corpus callosum obtained in this work. Heightened anatomical connectivity in the corpus callosum seems to correlate with higher intelligence. Moreover, IQ is intimately associated with diagnostic, due to the cognitive deficits associated with neuropsychiatric illnesses as SZ and BD [95]. Due to this, the effect of IQ on FA/MD was only assessed for descriptive proposes and not with the aim of controlling its effect in the main analyses of this work.

YE revealed significant positive effects on FA in superior corona radiata (R^2=0.084) and posterior thalamic radiation (R^2=0.059). This variable is strongly correlated with IQ (r = 0.405, p = 5.24×10^{-7}), due to this the results obtained were very similar. No effects of YE on MD survived the minimum cluster extent suggesting that this variable does not have an effect of MD.

4.2 Effect of diagnosis

The assessment of the effect of diagnosis on FA was performed for two different samples: (1) a sample with all subjects (n=143) and an (2) age-matched sample (n=115). As previously mentioned, the selection of an age-matched sample was conducted due to the significant effect of age in both FA and diagnosis. The effects of diagnosis on MD were acquired only for the whole-sample, because age did not show to have a significant effect on this measure.

Almost all effects of diagnosis reported are for an TFCE-uncorrected p-value < 0.01 (trends). Considering the whole-sample, no effect survived the minimum cluster extent for the main effect analysis. However, post-hoc analysis showed that REL are associated with larger values of when compared with HC in FA in genu of corpus callosum, at the significant level and with a large effect size (d = 0.787).

Besides significant effects, trends for this contrast were found in middle cerebellar peduncle, corpus callosum, cingulum, anterior thalamic radiation and inferior fronto-occipital fasciculus. The results suggest that REL have a higher intelligence than HC for the whole-sample, as high anatomical connectivity in the corpus callosum seems to correlate with a higher IQ. This fact is supported by the analysis of the demographics for the whole-sample that shows a higher mean of the IQ z-scores for REL (mean = 0.54) than HC (mean = -0.08).

In the post-hoc analysis for the whole-sample, the effects revealed that REL are assocated with greater FA values than SZ and BD. However, these effects did not reach significance. For the contrast REL > BD, trend effects were found in inferior longitudinal fasciculus, corpus callosum, inferior fronto-occipital fasciculus and cingulum. The most robust effect was located in inferior longitudinal fasciculus (d = 0.497). Moreover, for the contrast REL > SZ trends were evidenced in cingulum and inferior longitudinal fasciculus, being the strongest effect located in the inferior longitudinal fasciculus (d = 0.284). Lower values of FA in inferior longitudinal fasciculus for SZ and BD have been reported in literature described in previous sections of this work [68], [72], [75], supporting the results exhibit in this project. Moreover, the effects found in these studies were significant, suggesting that possibly the effects found for this study did not reach significance due to the small sample size.

For the age-matched subsample trend effects were found in the analysis of the main effect of diagnosis on FA in superior and inferior longitudinal fasciculus, cingulum, external capsule, inferior fronto-occipital fasciculus, corticospinal tract and corpus callosum. However, the effects sizes (η^2) were not robust and the effects did not reach significance. Post-hoc analyses revealed that REL have larger FA values when compared with SZ and BD. Significant effects were found for the contrast REL > BD in posterior corona radiata (d = 0.556). Trend effects for this contrast revealed effects in inferior cerebellar peduncle, corticopontine tract and cingulum. For the contrast REL > SZ only trend effects were found across the brain WM, more specifically in the cingulum, corticospinal tract, superior and inferior longitudinal fasciculus, uncinate fasciculus, corpus callosum and inferior fronto-occipital fasciculus. Furthermore, HC showed to have greater FA values than BD. Robust significant effects for this contrast were found in posterior corona radiata (d=0.749) and corticospinal tract (d=0.689). Trend effects were also uncovered in superior and inferior longitudinal fasciculus, corticospinal tract and cingulum.

For the age-matched sample, significant effects in the posterior corona radiata have been found for the contrast between BD and REL/HC, being the effect stronger when comparing BD with HC. Lower values of FA in the corona radiata of BD patients have been reported in a paper described in the literature review section of this work [73]. The paper previously enunciated, supports the findings of this work. Due to the association between corona radiata and the corticopontine tract, the corticobulbar tract, and the corticospinal tract it makes sense to have a significant effect for posterior corona radiata and corticospinal tract for the same contrast.

At last, the analysis of the influence of diagnosis on MD revealed that BD is associated with larger values of MD than REL and HC. Results showed that REL has reduced MD values when compared with BD in anterior thalamic radiation and inferior fronto-occipital fasciculus at the trend level. Significant effects were found for the contrast BD > HC in superior longitudinal fasciculus (d = 0.576) and anterior

thalamic radiation (d =0.086). However, the last significant effect enumerated was very weak. Greater values of MD for BD in the longitudinal fasciculus have been reported in a study described above, supporting the results found for this work [75]. Trends were found in inferior and superior longitudinal fasciculus, corpus callosum, inferior fronto-occipital fasciculus, posterior thalamic radiation.

As observed, the creation of an age-matched sample improved the results in comparison with the whole-sample. This statement is sustained by the fact that in the analysis of the effect of diagnosis on FA for the age-matched sample more effects survived the minimum cluster extent and reached significance. Nevertheless, the findings were similar for the analysis in both samples revealing reduced values of FA for SZ and BD when compared with REL.

A fairly number of studies has reported lower FA for SZ and BD when compared with healthy individuals [66]–[75], which sustains the effects revealed in this study. As previously stated reduced FA can be a result of a combination of factors that are related with the illness (SZ and BD). However, it is impossible to know which factors contribution are just based on the FA. Although it is not possible to clarify the exact cause of reductions in FA along the WM tracts in BD and SZ, several factors may contribute to the presence of disrupted tracts, such as alterations in axonal integrity, a reduced degree of myelination, a loss in axonal bundle coherence, or a variation of membrane permeability to water. Because WM is primarily composed of myelinated axons, the density of axons and myelin play a major role in DTI measurements of the diffusion indexes within WM tracts. The determination of MD helps to unravel which biological process is influencing the WM integrity. Increased MD in SZ and BD suggests reduced myelinization and reduced neural coherence as mentioned in [66], [69], [75]. The facts enunciated sustain the results obtained for the influence of diagnosis on MD.

Previous literature has associated brain function with WM microstructure integrity, by combining DTI and fMRI techniques. The work previously enunciated found a positive correlation between reward-related activation (mostly in the ventral striatum) with FA in WM regions as corpus callosum, inferior fronto-occipital fasciculus, cingulate bundle, anterior limb of internal capsule and anterior thalamic radiation [117]. Besides, previous work has also revealed that striatal disfunction plays a major role in SZ, as striatal dopamine dysfunction impacts on reward processing and learning and is present even at rest. Moreover, changes in intrinsic striatal activity occur in patients with SZ [118]. In this study, WM regions as corpus callosum and inferior fronto-occipital fasciculus revealed to be associated with lower values of FA in SZ and BD, as previously described in the results section. The regions previously enunciated seem to be associated with a decrease in ventral striatum activation and dysfunctions in striatal activation seem to be related with SZ. Therefore, the findings suggest that lower values of FA in corpus callosum and inferior-fronto occipital fasciculus may be related with dysfunctional activation in the striatum in SZ, and possibly in BD, as both disorders share genetic susceptibility and clinical symptoms.

4.3 Effect of PRS

The main effect of PRS on FA was very small and did not reach significance, across the brain. Positive and negative trend effects were found for the effect of PRS on FA. In effect, positive results were slightly larger than the negative ones which contradicts the main hypothesis of this work. Positive trend effects were found in uncinate fasciculus (R^2=0.048), corpus callosum (R^2=0.079), middle cerebellar peduncle (R^2=0.077) and anterior thalamic radiation (R^2=0.045). Negative trend effects were showed in superior longitudinal fasciculus (R^2=0.010), anterior (R^2=0.049) and posterior thalamic radiation (R^2=0.062), corticopontine tract (R^2=0.090), middle (R^2=0.018) and superior cerebellar peduncle (R^2=0.063), anterior corona radiata (R^2=0.054) and cingulum (R^2=0.053).

PRS did not significantly influence MD. Positive trend effects for the main effect of PRS on MD were found in inferior longitudinal fasciculus (R^2=0.053) and posterior corona radiata (R^2=0.067). Negative trend effects were evidenced in inferior cerebellar peduncle (R^2=0.061).

The results did not reveal what was hypothesized. As SZ PRS and reduced FA have been successfully associated with both SZ and BD, the expectation was that SZ PRS would be negatively correlated with FA. However, SZ PRS only explains about 9.4% of the variability between patients and controls [6]. This fact in combination with a small sample size in the diseased groups (n(SZ) =22; n(BD) =25) can be the reason why it is impossible to observe a significant effect of PRS on FA. In a previous study, with a larger sample size, that focused in the relation between SZ PRS and WM microstructure integrity, no significant effect of SZ PRS and FA/MD was found [77]. Nevertheless, the highly heritability of SZ and BD suggests that genetic factors pose a major contribution in brain alterations related with these disorders, but nevertheless, the genetic risk variants that contribute most to the disorder, or to WM alterations, have not been detected by the PGC.

Furthermore, studies described in the literature review section, have positively associated SZ PRS, calculated based on PGC variants, with negative symptoms as blunted affect and emotional withdrawal [47], [48]. This may be the reason why it is difficult to detect a significant influence of SZ PRS on WM integrity, as the changes in WM integrity may be associated with genetic variants more related with positive symptoms.

In sum, the results found suggest that WM microstructure integrity may not be closely linked to the SZ PRS calculated based on the variants detected by PGC.

4.4 PRS x diagnosis effects

The interaction analysis aimed to find diagnostic-dependent effects of PRS on FA/MD. As hypothesised, I expect that the effects of SZ PRS and FA/MD are different between the diagnostic groups.

Only effects at the trend level were found for the analysis of the interaction between PRS and diagnosis, both for FA and MD. Trends showed that the correlation between PRS and FA was larger for

BD than for HC in middle cerebellar peduncle, superior and inferior longitudinal fasciculus, cingulum, pontine crossing tract, corpus callosum, inferior fronto-occipital fasciculus, cortical spinal tract and uncinate fasciculus. For this contrast, the effects were weak, being the strongest one located in uncinate fasciculus (R^2=0.043). Moreover, trends showed that the correlation between PRS and FA was larger in REL than in HC in inferior fronto-occipital fasciculus (R^2=0.007). Trends also showed that the correlation between PRS and FA is higher in BD than in SZ in middle cerebellar peduncle (R^2=0.009) and corticopontine tract (R^2=0.019). At last, PRS showed to have larger correlation in REL than in SZ in middle cerebellar peduncle (R^2=0.015). For the age-matched sample only trend effects with less than 50 voxels were found as the effects of SZ PRS on WM integrity were even smaller when the sample was reduced (n = 115).

For the correlation of PRS and MD only one trend effect was found in inferior fronto-occipital fasciculus (R^2=0.011). This effect revealed that the correlation between PRS and MD is higher in HC than BD.

Based on the hypothesis, it was expected that the effects of SZ PRS on FA/MD are different between diagnostic groups, specially between diseased groups and controls. The results showed differences in the effects of SZ PRS on FA/MD between the different diagnostic groups, but only at the trend level. The effects found were also very weak suggesting that at least the present SZ PRS [6] and FA/MD are not strongly correlated.

One of the possible reasons for the results described is that the sample size may be too small for the study to have power to find a correlation between SZ PRS and FA. However, as mentioned before, no effect of SZ PRS on WM microstructure was found in a contemporaneous study that comprised a large sample size [77]. As such, I could not disprove the null hypothesis of this study. An explanation other than the lack of power is that SZ PRS, calculated based on PGC, does not detect SZ PRS diagnosis-dependent effects on WM microstructure. For example, if SZ PRS, calculated by PGC, is more associated with negative symptoms [47], [48] it may be that I did not detect an effect of the SZ PRS on WM microstructure due to this brain phenotype being possibly more associated with positive symptoms. Nevertheless, these are tentive explanations that warrant further testing.

5 Conclusions and future work

5.1 Conclusions

Based on the results found in this study it is possible to conclude that the influence of the different diagnostic groups on FA and MD is distinct. As previously mentioned, lower values of FA are associated with diseased groups (SZ, BD), suggesting alterations in axonal integrity, a reduced degree of myelination in SZ and BD. As MD and FA have antagonistic behaviour, high MD seems to be related with the presence of SZ and BD. The acquisition of both diffusion measures helps understanding the biological mechanisms that lead to impairments in WM integrity, as MD represents the magnitude of diffusion and FA the directionality of the diffusion.

Regarding PRS, no significant correlations between this score and the diffusion measures were found, both with and without considering diagnostic groups. Taking in account the results found for this work and in previous literature, SZ PRS, calculated based on the variants detected by PGC, seems not to be intimately linked with WM microstructure integrity. Although the sample size used in this work was small, a previous study has used a sample about 5 times larger and also found no correlation between the SZ PRS and WM microstructure integrity [77].

5.2 Limitations and future work

Several limitations can be pointed out in this work. The biggest limitation of this work was the sample size in the diseased groups that revealed to be too small to detect alterations in the brain resultant of genetic variations between patients and healthy subjects. As SZ PRS only explains 9.4% of the variance between patients and controls in a Caucasian sample [6], a better SZ PRS (i.e that explains a higher percentage of variance) or a bigger sample size would be necessary to take conclusions about the relationship between SZ PRS and FA/MD in SZ and BD when compared with healthy subjects. Future projects should focus on conducting this study with a larger sample size and/or a better SZ PRS. Furthermore, the PRS can also be calculated with variants related with more than one disorder, in this case the ideal would be to calculate a PRS based both on SZ and BD variants. Previous studies have calculated cross-disorder PRS based on several psychiatric illnesses [119].

The treatment with different antipsychotics and antidepressants and their dosage might also affect the brain structure and consequently the integrity of the WM. Studies have reported that WM volume losses

may be associated with the intensity of the antipsychotic treatment [120]. Ideally, the medication and its dosage should be controlled by selecting participants that receive the same treatment. A different approach would be to include CPZ-equivalent in the analysis design. This standardized quantitative method is used to compare medication doses between patients with different medication [121].

The results presented in this work can only be validated by performing the same project in a larger independent sample.

Bibliography

[1] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders : DSM-5*. Arlington, VA: American Psychiatric Publishing, 2013, ISBN: 9780890425541 089042554X 9780890425558 0890425558.

[2] R. Gurung and D. P. Prata, "What is the impact of genome-wide supported risk variants for schizophrenia and bipolar disorder on brain structure and function? A systematic review", eng, *Psychol Med*, vol. 45, no. 12, pp. 2461–2480, 2015.

[3] N. D. Woodward, "The course of neuropsychological impairment and brain structure abnormalities in psychotic disorders", *Neurosci. Res.*, vol. 102, pp. 39–46, 2016, ISSN: 0168-0102.

[4] N. Craddock and M. J. Owen, "The beginning of the end for the Kraepelinian dichotomy", *Br. J. Psychiatry*, vol. 186, no. 5, pp. 364–366, 2005.

[5] Psychiatric Genomics Consortium Schizophrenia Working Group, "Biological insights from 108 schizophrenia-associated genetic loci", eng, *Nature*, vol. 511, no. 7510, pp. 421–427, 2014, ISSN: 14764687. arXiv: *NIHMS150003*.

[6] E. Vassos, M. Di Forti, J. Coleman, *et al.*, "An Examination of Polygenic Score Risk Prediction in Individuals With First-Episode Psychosis", eng, *Biol Psychiatry*, vol. 81, no. 6, pp. 470–477, 2017.

[7] E. J. Mallas, F Carletti, C. A. Chaddock, *et al.*, "Genome-wide discovered psychosis-risk gene ZNF804A impacts on white matter microstructure in health, schizophrenia and bipolar disorder", eng, *PeerJ*, vol. 4, e1570, 2016.

[8] L. J. O'Donnell and C. F. Westin, "An introduction to diffusion tensor image analysis", eng, *Neurosurg Clin N Am*, vol. 22, no. 2, pp. 185–96, viii, 2011.

[9] M. S. Keshavan, D. W. Morris, J. A. Sweeney, *et al.*, "A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale", eng, *Schizophr Res*, vol. 133, no. 1-3, pp. 250–254, 2011.

[10] R. Freedman, "Schizophrenia", eng, *N Engl J Med*, vol. 349, no. 18, pp. 1738–1749, 2003, ISSN: 0028-4793.

[11] D. I. Driver, N. Gogtay, and J. L. Rapoport, "Childhood onset schizophrenia and early onset schizophrenia spectrum disorders.", eng, *Child Adolesc. Psychiatr. Clin. N. Am.*, vol. 22, no. 4, pp. 539–555, 2013, ISSN: 1558-0490 (Electronic).

[12] H. Hafner, K. Maurer, W. Loffler, *et al.*, "The epidemiology of early schizophrenia. Influence of age and gender on onset and early course", eng, *Br J Psychiatry Suppl*, no. 23, pp. 29–38, 1994.

[13] I. Grande, M. Berk, B. Birmaher, *et al.*, "Bipolar disorder", eng, *Lancet*, vol. 387, no. 10027, pp. 1561–1572, 2016.

[14] E. Paykel, "Manic–Depressive Illness: Bipolar Disorders and Recurrent Depression (2nd edn)", *Br. J. Psychiatry*, vol. 193, no. 1, pp. 86–87, 2008.

[15] K. E. Muñoz, L. W. Hyde, and A. R. Hariri, "Imaging Genetics", eng, *J Am Acad Child Adolesc Psychiatry*, vol. 48, no. 4, pp. 356–361, 2009.

[16] A. Tan, G. R. Abecasis, and H. M. Kang, "Unified representation of genetic variants", *Bioinformatics*, vol. 31, no. 13, pp. 2202–2204, 2015.

[17] M. G. Henriksen, J Nordgaard, and L. B. Jansson, "Genetics of Schizophrenia: Overview of Methods, Findings and Limitations", eng, *Front Hum Neurosci*, vol. 11, p. 322, 2017.

[18] Psychiatric GWAS Consortium Bipolar Disorder Working Group, "Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4", eng, *Nat. Genet.*, vol. 43, no. 10, pp. 977–985, 2011, ISSN: 10614036.

[19] R. Maier, G. Moser, G. B. Chen, *et al.*, "Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder", eng, *Am J Hum Genet*, vol. 96, no. 2, pp. 283–294, 2015.

[20] N. Chatterjee, B. Wheeler, J. Sampson, *et al.*, "Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies", eng, *Nat Genet*, vol. 45, no. 4, 400–5, 405e1–3, 2013.

[21] N. R. Wray, S. H. Lee, D Mehta, *et al.*, "Research review: Polygenic methods and their application to psychiatric traits", eng, *J Child Psychol Psychiatry*, vol. 55, no. 10, pp. 1068–1087, 2014.

[22] S. M. Purcell, N. R. Wray, J. L. Stone, *et al.*, "Common polygenic variation contributes to risk of schizophrenia and bipolar disorder", eng, *Nature*, vol. 460, no. 7256, pp. 748–752, 2009.

[23] G. A. Preston and D. R. Weinberger, "Intermediate phenotypes in schizophrenia: a selective review", *Dialogues Clin. Neurosci.*, vol. 7, no. 2, pp. 165–179, 2005, ISSN: 1294-8322 (Print).

[24] K. W. Lee, P. S. Woon, Y. Y. Teo, *et al.*, "Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt?", eng, *Neurosci Biobehav Rev*, vol. 36, no. 1, pp. 556–571, 2012.

[25] C. C. Schultz, T. W. Mühleisen, I Nenadic, *et al.*, "Common variation in NCAN, a risk factor for bipolar disorder and schizophrenia, influences local cortical folding in schizophrenia", *Psychol Med*, vol. 44, no. 4, pp. 811–820, 2013.

[26] S. Bhat, D. T. Dao, C. E. Terrillion, *et al.*, "CACNA1C (Ca(v)1.2) in the pathophysiology of psychiatric disease", eng, *Prog Neurobiol*, vol. 99, no. 1, pp. 1–14, 2012.

[27] H. Stefansson, R. A. Ophoff, S. Steinberg, *et al.*, "Common variants conferring risk of schizophrenia", *Nature*, vol. 460, p. 744, 2009.

[28] Z. Iqbal, G. Vandeweyer, M. van der Voet, *et al.*, "Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders", *Hum. Mol. Genet.*, vol. 22, no. 10, pp. 1960–1970, 2013.

[29] S. Kittel-Schneider, T. Wobrock, H. Scherk, *et al.*, "Influence of DGKH variants on amygdala volume in patients with bipolar affective disorder and schizophrenia", eng, *Eur. Arch. Psychiatry Clin. Neurosci.*, vol. 265, no. 2, pp. 127–136, 2015, ISSN: 14338491.

[30] H. Raum, B. Dietsche, A. Nagels, *et al.*, "A genome-wide supported psychiatric risk variant in NCAN influences brain function and cognitive performance in healthy subjects", eng, *Hum Brain Mapp*, vol. 36, no. 1, pp. 378–390, 2015.

[31] E. G. Porter and E. C. Dykhuizen, "Individual Bromodomains of Polybromo-1 Contribute to Chromatin Association and Tumor Suppression in Clear Cell Renal Carcinoma", eng, *J Biol Chem*, vol. 292, no. 7, pp. 2601–2610, 2017.

[32] S. Groeschel, G. E. Hagberg, T. Schultz, *et al.*, "Assessing White Matter Microstructure in Brain Regions with Different Myelin Architecture Using MRI", eng, *PLoS One*, vol. 11, no. 11, e0167274, 2016.

[33] M. H. A. Bopp, R Zollner, A Jansen, *et al.*, "White matter integrity and symptom dimensions of schizophrenia: A diffusion tensor imaging study", eng, *Schizophr Res*, vol. 184, pp. 59–68, 2017.

[34] M. A. Brown and R. C. Semelka, *MRI: Basic Principles and Applications*. Wiley, 2011, ISBN: 9780470920862.

[35] N. B. Smith and A Webb, *Introduction to Medical Imaging: Physics, Engineering and Clinical Applications*. Cambridge University Press, 2010, ISBN: 9781139492041.

[36] P. J. Basser, J Mattiello, and D Lebihan, "Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo", *J. Magn. Reson. Ser. B*, vol. 103, no. 3, pp. 247–254, 1994.

[37] P. J. Basser, J Mattiello, and D LeBihan, "MR diffusion tensor spectroscopy and imaging", eng, *Biophys J*, vol. 66, no. 1, pp. 259–267, 1994.

[38] A. L. Alexander, J. E. Lee, M Lazar, *et al.*, "Diffusion Tensor Imaging of the Brain", eng, *Neurotherapeutics*, vol. 4, no. 3, pp. 316–329, 2007.

[39] P. Hagmann, L. Jonasson, P. Maeder, *et al.*, "Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond", eng, *Radiographics*, vol. 26 Suppl 1, S205–23, 2006.

[40] M. Tesli, T. Espeseth, F. Bettella, *et al.*, "Polygenic risk score and the psychosis continuum model", eng, *Acta Psychiatr Scand*, vol. 130, no. 4, pp. 311–317, 2014.

[41] D. M. Ruderfer, A. H. Fanous, S Ripke, *et al.*, "Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia", eng, *Mol Psychiatry*, vol. 19, no. 9, pp. 1017–1024, 2014.

[42] J. W. Smoller, K. Kendler, N. Craddock, *et al.*, "Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis", eng, *Lancet*, vol. 381, no. 9875, pp. 1371–1379, 2013, ISSN: 1474547X. arXiv: *NIHMS150003*.

[43] Y. Milaneschi, F. Lamers, W. J. Peyrot, *et al.*, "Polygenic dissection of major depression clinical heterogeneity", eng, *Mol Psychiatry*, vol. 21, no. 4, pp. 516–522, 2016.

[44] S. Stringer, R. S. Kahn, L. D. de Witte, *et al.*, "Genetic liability for schizophrenia predicts risk of immune disorders", eng, *Schizophr Res*, vol. 159, no. 2-3, pp. 347–352, 2014.

[45] C. M. Nievergelt, A. X. Maihofer, M Mustapic, *et al.*, "Genomic predictors of combat stress vulnerability and resilience in U.S. Marines: A genome-wide association study across multiple ancestries implicates PRTFDC1 as a potential PTSD gene", eng, *Psychoneuroendocrinology*, vol. 51, pp. 459–471, 2015.

[46] E. M. Derks, J. A. Vorstman, S Ripke, *et al.*, "Investigation of the genetic association between quantitative measures of psychosis and schizophrenia: a polygenic risk score analysis", eng, *PLoS One*, vol. 7, no. 6, e37852, 2012.

[47] A. H. Fanous, B Zhou, S. H. Aggen, *et al.*, "Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms", eng, *Am J Psychiatry*, vol. 169, no. 12, pp. 1309–1317, 2012.

[48] H. J. Jones, E Stergiakouli, K. E. Tansey, *et al.*, "Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population", eng, *JAMA Psychiatry*, vol. 73, no. 3, pp. 221–228, 2016.

[49] D. Sieradzka, R. A. Power, D. Freeman, *et al.*, "Are genetic risk factors for psychosis also associated with dimension-specific psychotic experiences in adolescence?", eng, *PLoS One*, vol. 9, no. 4, e94398, 2014.

[50] T. Lencz, E. Knowles, G. Davies, *et al.*, "Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT)", eng, *Mol Psychiatry*, vol. 19, no. 2, pp. 168–174, 2014.

[51] L. Hubbard, K. E. Tansey, D. Rai, *et al.*, "Evidence of Common Genetic Overlap Between Schizophrenia and Cognition", eng, *Schizophr Bull*, vol. 42, no. 3, pp. 832–842, 2016.

[52] A. M. McIntosh, A Gow, M Luciano, *et al.*, "Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age", eng, *Biol Psychiatry*, vol. 73, no. 10, pp. 938–943, 2013.

[53] A. Hatzimanolis, P. Bhatnagar, A. Moes, *et al.*, "Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood", eng, *Am J Med Genet B Neuropsychiatr Genet*, vol. 168b, no. 5, pp. 392–401, 2015.

[54] R. Shafee, P. Nanda, J. L. Padmanabhan, *et al.*, "Polygenic risk for schizophrenia and measured domains of cognition in individuals with psychosis and controls.", eng, *Transl. Psychiatry*, vol. 8, no. 1, p. 78, 2018, ISSN: 2158-3188 (Electronic).

[55] E. Walton, J. Turner, R. L. Gollub, *et al.*, "Cumulative genetic risk and prefrontal activity in patients with schizophrenia", eng, *Schizophr Bull*, vol. 39, no. 3, pp. 703–711, 2013.

[56] E. Walton, D. Geisler, P. H. Lee, *et al.*, "Prefrontal inefficiency is associated with polygenic risk for schizophrenia", eng, *Schizophr Bull*, vol. 40, no. 6, pp. 1263–1271, 2014.

[57] K. Kauppi, L. T. Westlye, M. Tesli, *et al.*, "Polygenic risk for schizophrenia associated with working memory-related prefrontal brain activation in patients with schizophrenia and healthy controls", eng, *Schizophr Bull*, vol. 41, no. 3, pp. 736–743, 2015.

[58] T. M. Lancaster, N Ihssen, L. M. Brindley, *et al.*, "Associations between polygenic risk for schizophrenia and brain function during probabilistic learning in healthy individuals", eng, *Hum Brain Mapp*, vol. 37, no. 2, pp. 491–500, 2016.

[59] D. Cosgrove, D. Harold, O. Mothersill, *et al.*, "MiR-137-derived polygenic risk: effects on cognitive performance in patients with schizophrenia and controls", eng, *Transl Psychiatry*, vol. 7, no. 1, e1012, 2017.

[60] J. A. Miller, M. A. Scult, E. D. Conley, *et al.*, "Effects of Schizophrenia Polygenic Risk Scores on Brain Activity and Performance During Working Memory Subprocesses in Healthy Young Adults", eng, *Schizophr Bull*, 2017.

[61] A. F. Terwisscha van Scheltinga, S. C. Bakker, N. E. van Haren, *et al.*, "Genetic schizophrenia risk variants jointly modulate total brain and white matter volume", eng, *Biol Psychiatry*, vol. 73, no. 6, pp. 525–531, 2013.

[62] V. Oertel-Knochel, T. M. Lancaster, C. Knochel, *et al.*, "Schizophrenia risk variants modulate white matter volume across the psychosis spectrum: evidence from two independent cohorts", eng, *Neuroimage Clin*, vol. 7, pp. 764–770, 2015.

[63] F. Harrisberger, R. Smieskova, C. Vogler, *et al.*, "Impact of polygenic schizophrenia-related risk and hippocampal volumes on the onset of psychosis", eng, *Transl Psychiatry*, vol. 6, no. 8, e868, 2016.

[64] A. N. Voineskos, D. Felsky, A. L. Wheeler, *et al.*, "Limited Evidence for Association of Genome-Wide Schizophrenia Risk Variants on Cortical Neuroimaging Phenotypes", eng, *Schizophr Bull*, vol. 42, no. 4, pp. 1027–1036, 2016.

[65] E. Neilson, C. Bois, J. Gibson, *et al.*, "Effects of environmental risks and polygenic loading for schizophrenia on cortical thickness", eng, *Schizophr Res*, vol. 184, pp. 128–136, 2017.

[66] C. Zhuo, X. Ma, H. Qu, *et al.*, "Schizophrenia Patients Demonstrate Both Inter-Voxel Level and Intra-Voxel Level White Matter Alterations", eng, *PLoS One*, vol. 11, no. 9, e0162656, 2016.

[67] P. V. Viher, K Stegmayer, S Giezendanner, *et al.*, "Cerebral white matter structure is associated with DSM-5 schizophrenia symptom dimensions", eng, *Neuroimage Clin*, vol. 12, pp. 93–99, 2016.

[68] T. A. Hummer, M. M. Francis, J. L. Vohs, *et al.*, "Characterization of white matter abnormalities in early-stage schizophrenia", eng, *Early Interv Psychiatry*, 2016.

[69] R. A. Kanaan, M. M. Picchioni, C McDonald, *et al.*, "White matter deficits in schizophrenia are global and don't progress with age", eng, *Aust N Z J Psychiatry*, vol. 51, no. 10, pp. 1020–1031, 2017.

[70] K. Subramaniam, J. Gill, M. Fisher, *et al.*, "White matter microstructure predicts cognitive training-induced improvements in attention and executive functioning in schizophrenia", eng, *Schizophr Res*, 2017.

[71] F. Lin, S. Weng, B. Xie, *et al.*, "Abnormal frontal cortex white matter connections in bipolar disorder: a DTI tractography study", eng, *J Affect Disord*, vol. 131, no. 1-3, pp. 299–306, 2011.

[72] E. Ambrosi, M. C. Rossi-Espagnet, G. D. Kotzalidis, *et al.*, "Structural brain alterations in bipolar disorder II: a combined voxel-based morphometry (VBM) and diffusion tensor imaging (DTI) study", eng, *J Affect Disord*, vol. 150, no. 2, pp. 610–615, 2013.

[73] A. Ji, D. Godwin, J. Rutlin, *et al.*, "Tract-based analysis of white matter integrity in psychotic and nonpsychotic bipolar disorder", eng, *J Affect Disord*, vol. 209, pp. 124–134, 2017.

[74] A. Mahapatra, S. K. Khandelwal, P. Sharan, *et al.*, "Diffusion tensor imaging tractography study in bipolar disorder patients compared to first-degree relatives and healthy controls", eng, *Psychiatry Clin Neurosci*, vol. 71, no. 10, pp. 706–715, 2017.

[75] L. Squarcina, M. Bellani, M. G. Rossetti, *et al.*, "Similar white matter changes in schizophrenia and bipolar disorder: A tract-based spatial statistics study.", eng, *PLoS One*, vol. 12, no. 6, e0178089, 2017, ISSN: 1932-6203 (Electronic).

[76] H. C. Whalley, E. Sprooten, S. Hackett, *et al.*, "Polygenic risk and white matter integrity in individuals at high risk of mood disorder.", eng, *Biol. Psychiatry*, vol. 74, no. 4, pp. 280–286, 2013, ISSN: 1873-2402 (Electronic).

[77] L. M. Reus, X Shen, J Gibson, *et al.*, "Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in UK Biobank.", eng, *Sci. Rep.*, vol. 7, p. 42 140, 2017, ISSN: 2045-2322 (Electronic).

[78] M. P. Allin, D Kontis, M Walshe, *et al.*, "White matter and cognition in adults who were born preterm", eng, *PLoS One*, vol. 6, no. 10, e24525, 2011.

[79] C. A. Chaddock, G. J. Barker, N Marshall, *et al.*, "White matter microstructural impairments and genetic liability to familial bipolar I disorder", eng, *Br J Psychiatry*, vol. 194, no. 6, pp. 527–534, 2009.

[80] R. A. Kanaan, C Chaddock, M Allin, *et al.*, "Gender influence on white matter microstructure: a tract-based spatial statistics analysis", eng, *PLoS One*, vol. 9, no. 3, e91109, 2014.

[81] M. Kyriakopoulos, R. Perez-Iglesias, J. B. Woolley, *et al.*, "Effect of age at onset of schizophrenia on white matter abnormalities", eng, *Br J Psychiatry*, vol. 195, no. 4, pp. 346–353, 2009.

[82] M. M. Picchioni, T Toulopoulou, S Landau, *et al.*, "Neurological abnormalities in schizophrenic twins", eng, *Biol Psychiatry*, vol. 59, no. 4, pp. 341–348, 2006.

[83] S. S. Shergill, R. A. Kanaan, X. A. Chitnis, *et al.*, "A diffusion tensor imaging study of fasciculi in schizophrenia", eng, *Am J Psychiatry*, vol. 164, no. 3, pp. 467–473, 2007.

[84] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition: DSM-IV-TR®*. American Psychiatric Publishing, 2000, ISBN: 9780890420256.

[85] IBM Corp., *IBM SPSS Statistics 24.0*. Armonk, NY, 2016.

[86] D. Wechsler, *Wechsler Adult Intelligence Scale-Third Edition*. Psychological Corporation, 1997.

[87] D. Wechsler, *Wechsler adult intelligence scale-revised*. Psychological Corporation, 1981.

[88] D. Wechsler, *Wechsler Adult Intelligence Scale-Fourth Edition*. Psychological Corporation, 2008.

[89] J. Venegas and E. Clark, *National Adult Reading Test*. 2011, p. 1705, ISBN: 978-0-387-79947-6.

[90] D. J. Madden, I. J. Bennett, A. Burzynska, *et al.*, "Diffusion tensor imaging of cerebral white matter integrity in cognitive aging.", eng, *Biochim. Biophys. Acta*, vol. 1822, no. 3, pp. 386–400, 2012, ISSN: 0006-3002 (Print).

[91] D. H. Salat, "The declining infrastructure of the aging brain.", eng, *Brain Connect.*, vol. 1, no. 4, pp. 279–293, 2011, ISSN: 2158-0022 (Electronic).

[92] O. Carmichael and S. Lockhart, "The role of diffusion tensor imaging in the study of cognitive aging.", eng, *Curr. Top. Behav. Neurosci.*, vol. 11, pp. 289–320, 2012, ISSN: 1866-3370 (Print).

[93] I. J. Bennett and D. J. Madden, "Disconnected aging: cerebral white matter integrity and age-related differences in cognition.", eng, *Neuroscience*, vol. 276, pp. 187–205, 2014, ISSN: 1873-7544 (Electronic).

[94] F. Kamangar, "Confounding variables in epidemiologic studies: basics and beyond.", eng, *Arch. Iran. Med.*, vol. 15, no. 8, pp. 508–516, 2012, ISSN: 1735-3947 (Electronic).

[95] F. J. Navas-Sanchez, Y. Aleman-Gomez, J. Sanchez-Gonzalez, *et al.*, "White matter microstructure correlates of mathematical giftedness and intelligence quotient.", eng, *Hum. Brain Mapp.*, vol. 35, no. 6, pp. 2619–2631, 2014, ISSN: 1097-0193 (Electronic).

[96] I. Illumina, *GenomeStudio Software 2011.1*. San Diego: Illumina, 2011.

[97] S. Purcell and C. Chang, *PLINK (1.9)*.

[98] S. Purcell, B. Neale, K. Todd-Brown, *et al.*, "PLINK: a tool set for whole-genome association and population-based linkage analyses.", eng, *Am. J. Hum. Genet.*, vol. 81, no. 3, pp. 559–575, 2007, ISSN: 0002-9297 (Print).

[99] J. Euesden, C. M. Lewis, and P. F. O. Reilly, *Genome analysis PRSice : Polygenic Risk Score software*, December 2014. Bioinformatics, 2015, vol. 31, pp. 1466–1468.

[100] A. L. Price, N. J. Patterson, R. M. Plenge, *et al.*, "Principal components analysis corrects for stratification in genome-wide association studies", *Nat. Genet.*, vol. 38, p. 904, 2006.

[101] D. K. Jones, S. C. Williams, D Gasston, *et al.*, "Isotropic resolution diffusion tensor imaging with whole brain acquisition in a clinically acceptable time", eng, *Hum Brain Mapp*, vol. 15, no. 4, pp. 216–230, 2002.

[102] M. Jenkinson, C. F. Beckmann, T. E. Behrens, *et al.*, "FSL", eng, *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012, ISSN: 1053-8119.

[103] J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging", eng, *Neuroimage*, vol. 125, pp. 1063–1078, 2016.

[104] S. M. Smith, "Fast robust automated brain extraction", eng, *Hum Brain Mapp*, vol. 17, no. 3, pp. 143–155, 2002.

[105] T. E. Behrens, M. W. Woolrich, M Jenkinson, *et al.*, "Characterization and propagation of uncertainty in diffusion-weighted MR imaging", eng, *Magn Reson Med*, vol. 50, no. 5, pp. 1077–1088, 2003.

[106] S. M. Smith, M Jenkinson, H Johansen-Berg, *et al.*, "Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data", eng, *Neuroimage*, vol. 31, no. 4, pp. 1487–1505, 2006.

[107] K. J. Friston, "Models of brain function in neuroimaging", eng, *Annu Rev Psychol*, vol. 56, pp. 57–87, 2005.

[108] A. M. Winkler, G. R. Ridgway, M. A. Webster, *et al.*, "Permutation inference for the general linear model", eng, *Neuroimage*, vol. 92, pp. 381–397, 2014.

[109] S. M. Smith and T. E. Nichols, "Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference.", eng, *Neuroimage*, vol. 44, no. 1, pp. 83–98, 2009, ISSN: 1095-9572 (Electronic).

[110] H. Cheng, S. D. Newman, J. S. Kent, *et al.*, "White matter abnormalities of microstructure and physiological noise in schizophrenia.", eng, *Brain Imaging Behav.*, vol. 9, no. 4, pp. 868–877, 2015, ISSN: 1931-7565 (Electronic).

[111] S. Poletti, E. Mazza, I. Bollettini, *et al.*, "Adverse childhood experiences influence white matter microstructure in patients with schizophrenia.", eng, *Psychiatry Res.*, vol. 234, no. 1, pp. 35–43, 2015, ISSN: 1872-7123 (Electronic).

[112] S. Giezendanner, S. Walther, N. Razavi, *et al.*, "Alterations of white matter integrity related to the season of birth in schizophrenia: a DTI study.", eng, *PLoS One*, vol. 8, no. 9, e75508, 2013, ISSN: 1932-6203 (Electronic).

[113] S. Mori, S. Wakana, P. van Zijl, *et al.*, *MRI Atlas of Human White Matter*, 1st. Elsevier Science, 2005, p. 276, ISBN: 9780080456164.

[114] J. A. Durlak, "How to Select, Calculate, and Interpret Effect Sizes", *J. Pediatr. Psychol.*, vol. 34, no. 9, pp. 917–928, 2009, ISSN: 0146-8693.

[115] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd. New York: Routledge, 1988, p. 567, ISBN: 9781134742707.

[116] H. Takao, N. Hayashi, and K. Ohtomo, "Sex dimorphism in the white matter: fractional anisotropy and brain size.", eng, *J. Magn. Reson. Imaging*, vol. 39, no. 4, pp. 917–923, 2014, ISSN: 1522-2586 (Electronic).

[117] K. Koch, G. Wagner, C. Schachtzabel, *et al.*, "Association between white matter fiber structure and reward-related reactivity of the ventral striatum.", *Hum. Brain Mapp.*, vol. 35, no. 4, pp. 1469–1476, 2014, ISSN: 1097-0193 (Electronic).

[118] C. Sorg, A. Manoliu, S. Neufang, *et al.*, "Increased Intrinsic Brain Activity in the Striatum Reflects Symptom Dimensions in Schizophrenia", *Schizophr. Bull.*, vol. 39, no. 2, pp. 387–395, 2013, ISSN: 0586-7614 (Print).

[119] T. Wang, X. Zhang, A. Li, *et al.*, "Polygenic risk for five psychiatric disorders and cross-disorder and disorder-specific neural connectivity in two independent populations.", eng, *NeuroImage. Clin.*, vol. 14, pp. 441–449, 2017, ISSN: 2213-1582 (Electronic).

[120] T. A. Tishler, G. Bartzokis, P. H. Lu, *et al.*, "Abnormal Trajectory of Intracortical Myelination in Schizophrenia Implicates White Matter in Disease Pathophysiology and the Therapeutic Mechanism of Action of Antipsychotics.", eng, *Biol. psychiatry. Cogn. Neurosci. neuroimaging*, vol. 3, no. 5, pp. 454–462, 2018, ISSN: 2451-9030 (Electronic).

[121] N. C. Andreasen, M. Pressler, P. Nopoulos, *et al.*, "Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs.", eng, *Biol. Psychiatry*, vol. 67, no. 3, pp. 255–262, 2010, ISSN: 1873-2402 (Electronic).

A Apendix

A.1 Johns Hopkins University ICBM-DTI-81 white-matter labels atlas

This atlas is based on probabilistic tensor maps obtained from 81 normal subjects acquired under an initiative of the International Consortium of Brain Mapping. The subjects were normal right-handed adults ranging from 18 to 59 years of age.

All studies were obtained on 1.5T MR units (Siemens). DTI data was acquired by using a single-shot, echo-planar imaging sequence with sensitivity encoding and a parallel imaging factor of 2.0. Transverse sections of 2.5 mm thickness were acquired parallel to the anterior commissure-posterior commissure line.

The raw diffusion-weighted images were first co-registered to one of the least diffusion-weighted images and corrected for subject motion with 6-mode rigid transformation with Automated Image Registgration.

The average of all diffusion-weighted images was calculated and used for a DTI-based anatomic image. For anatomical images to drive the normalization process, all diffusion images were used. These images were normalized to the template (ICBM-152) using a 12-mode affine or 4th order polynomial non-linear transformation of AIR. The transformation matrix was then applied to the calculated diffusion tensor field. In the WM parcellation map, deep WM regions were manually segmented into various anatomic structures based on fiber orientation information. In the WM parcelation map, the following WM structures are identified and partitioned:

Tracts in the brainstem:

Corticospinal tracts

Medial lemniscus

Medial longitudinal fasciculus

Inferior cerebellar peduncle

Middle cerebellar peduncle

Superior cerebellar peduncle

Projection fibers:

Corona radiata

Anterior limb of internal capsule

Posterior limb of internal capsule

Retrolenticular part of the internal capsule

Cerebral peduncle

Association fibers:

Superior longitudinal fasciculus

Superior fronto-occipital fasciculus

Uncinate fasciculus

Inferior fronto-occipital fasciculus / Uncinate fasciculus

Inferior fronto-occipital fasciculus / Inferior longitudinal fasciculus

Sagittal Stratum

External capsule

Cingulum

Fornix and Stria terminalis

Commissural fibers:

Anterior commissure

Corpus callosum

Tapetum

A.2 MRI Atlas of the human white matter

Images were acquired with a 1.5 T Philips GyroscanNT scanner. Diffusion data were acquired using a single-shot EPI sequence with parallel imaging. The imaging matrix was 112×112 with a field of view of 246×246 mm (nominal resolution of 2.2 mm), which was zerofilled to a 256×256 matrix. The image orientation was axial with 2.2 mm slice thickness, which was aligned parallel to the anterior–posterior commissure line. A total of 55 slices covered the entire cerebral hemispheres and the brainstem. The diffusion weighting was encoded along 30 independent orientations with maximum b = 700 mm2/s. Five additional images with minimal diffusion weighting were also acquired. A co-registered MPRAGE image (T1-weighted image) with the same resolution was also recorded for anatomical guidance.

The 3D tract reconstruction was performed using the FACT method, which performs a straightforward linear line propagation based on the v1 vector angle. For fiber reconstruction, a tract of interest first needs to be identified and marked by a region of interest.

The nomenclature of the white matter structures included in this atlas is presented below.

Tracts in the brainstem:

superior cerebellar peduncle

middle cerebellar peduncle

inferior cerebellar peduncle

corticospinal tract

medial lemniscus

Projection Fibers:

superior thalamic radiation

anterior thalamic radiation

posterior thalamic radiation

corticopontine tract

corticospinal tract

corticobulbar tract

corticoreticular tract

Association Fibers:

superior longitudinal fasciculus

inferior longitudinal fasciculus

superior frontooccipital fasciculus

inferior fronto-occipital fasciculus

uncinate fasciculus

cingulum

fornix

stria terminalis

Commissural fibers:

corpus callosum

B Apendix

The regions of the PRS x diagnosis interaction effects on FA/MD with less than 50 voxels are presented in Tables B.1, B.2, B.3, B.4, B.8, B.7. Each effect is characterised by the cluster extent(k), the 1-(p-value) and the peak MNI coordinates.

TABLE B.1: TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in SZ >PRS in HC				
2	0.993	51	52	69
2	0.998	137	110	94
PRS in BD >PRS in HC				
48	0.994	74	40	67
45	0.994	131	126	39
42	0.994	134	89	80
35	0.993	33	97	62
33	0.994	102	62	123
33	0.993	119	61	37
28	0.993	121	73	59
24	0.995	129	70	39
23	0.993	65	35	67
20	0.997	76	146	66
19	0.992	107	80	131
19	0.996	48	64	65
18	0.995	69	35	84
18	0.993	98	60	105
18	0.993	82	90	43
17	0.995	66	147	109
16	0.993	41	133	80
15	0.992	124	69	35
15	0.996	129	85	55
14	0.991	107	38	83
14	0.995	106	140	64
14	0.992	99	127	63
14	0.996	105	33	85
14	0.994	102	71	126
13	0.994	57	175	79
13	0.994	62	121	110
13	0.995	72	31	73
12	0.994	106	90	140
12	0.993	50	74	56
12	0.994	95	96	59
12	0.992	82	130	135

TABLE B.2: (Continuation) TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in BD >PRS in HC				
12	0.994	57	169	93
12	0.993	103	61	50
11	0.993	132	130	47
10	0.993	121	163	99
10	0.993	134	81	101
10	0.995	104	154	61
8	0.991	34	112	72
8	0.993	131	115	36
8	0.996	107	57	42
8	0.991	101	73	129
7	0.993	56	172	88
7	0.994	64	129	113
7	0.993	59	167	94
7	0.993	37	90	64
6	0.995	66	169	71
6	0.992	68	33	73
6	0.995	130	63	37
6	0.994	117	68	46
6	0.991	107	109	85
5	0.992	57	163	98
5	0.993	100	138	64
5	0.994	114	63	119
5	0.996	118	127	89
5	0.993	99	41	66
4	0.992	50	82	55
4	0.998	123	95	70
4	0.99	53	137	42
4	0.992	61	132	88
4	0.993	102	41	66
4	0.996	66	67	28
4	0.993	93	101	56
3	0.991	58	132	38
3	0.993	41	109	47
3	0.991	95	103	50
3	0.992	101	74	125
3	0.994	137	110	94
3	0.991	104	87	139
3	0.993	86	99	54
3	0.992	41	101	97
3	0.99	83	96	57
3	0.992	113	104	68
3	0.993	102	67	122
3	0.991	64	38	67
3	0.993	103	55	120
2	0.992	61	128	90
2	0.992	135	75	104
2	0.991	60	177	65
2	0.99	110	125	86
2	0.99	51	66	59
2	0.993	97	56	105
2	0.991	102	98	82

TABLE B.3: (Continuation) TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in BD >PRS in HC				
2	0.991	66	129	117
2	0.992	119	122	46
2	0.99	112	149	66
2	0.99	111	100	78
2	0.999	116	86	108
2	0.992	81	136	63
2	0.991	98	157	125
2	0.993	64	161	101
2	0.992	72	35	69
2	0.992	95	114	71
2	0.991	40	110	114
2	0.99	124	153	99
1	0.99	103	86	136
1	0.99	99	158	123
1	0.991	123	101	120
1	0.991	104	64	124
1	0.991	116	93	101
1	0.99	116	91	101
1	0.991	107	66	124
1	0.99	106	67	119
1	0.99	98	155	124
1	0.99	108	71	118
1	0.99	84	102	126
1	0.99	65	88	105
1	0.99	82	101	126
1	0.991	82	97	126
1	0.99	84	98	125
1	0.991	41	83	111
1	0.99	108	67	116
1	0.996	132	88	112
1	0.99	108	67	114
1	0.991	105	74	130
1	0.99	100	123	63
1	0.99	95	103	63
1	0.99	34	95	63
1	0.991	32	95	63
1	0.991	72	181	60
1	0.99	104	146	60
1	0.991	95	101	58
1	0.99	51	104	51
1	0.99	86	99	50
1	0.99	51	99	50
1	0.99	100	60	50
1	0.99	115	93	136
1	0.99	51	105	49
1	0.99	107	57	49
1	0.991	110	58	48
1	0.99	97	86	45
1	0.991	129	129	43
1	0.991	114	69	43
1	0.99	124	56	43

TABLE B.4: (Continuation) TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in BD >PRS in HC				
1	0.991	74	90	39
1	0.991	77	67	35
1	0.99	126	74	34
1	0.99	67	69	32
1	0.992	86	96	31
1	0.991	97	58	50
1	0.99	137	80	99
1	0.991	119	162	97
1	0.991	54	161	97
1	0.99	55	163	95
1	0.99	124	166	88
1	0.991	108	36	84
1	0.99	69	34	77
1	0.99	70	31	77
1	0.99	98	107	76
1	0.991	30	110	75
1	0.991	83	104	74
1	0.991	105	127	63
1	0.99	95	120	71
1	0.99	136	77	71
1	0.99	68	35	70
1	0.991	87	98	69
1	0.99	98	97	68
1	0.992	102	44	68
1	0.99	108	147	66
1	0.991	81	138	66
1	0.99	101	39	66
1	0.991	103	138	65
1	0.99	127	77	65
1	0.99	104	125	64
1	0.99	104	110	72
1	0.991	105	74	130
1	0.99	100	123	63
1	0.99	95	103	63
PRS in REL >PRS in HC				
6	0.999	103	119	80
2	0.993	121	147	72
1	0.994	124	96	70
1	0.990	127	49	83
1	0.994	74	79	101
PRS in HC >PRS in SZ				
10	1.000	110	172	98
5	0.999	110	69	27
4	0.993	90	64	35
2	0.998	77	148	64
1	0.99	107	57	49
1	0.991	110	58	48
1	0.99	97	86	45
1	0.991	129	129	43
PRS in HC >PRS in REL				
3	0.993	64	34	90
1	0.991	77	46	109
1	0.993	63	83	102

TABLE B.5: (Continuation) TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in BD >PRS in SZ				
46	0.996	117	73	48
41	0.992	107	45	47
23	0.994	78	76	36
11	0.993	109	83	45
10	0.994	109	67	47
9	0.999	110	172	98
8	0.992	76	77	53
7	0.995	64	161	101
6	0.996	66	67	28
4	0.994	100	109	62
4	0.993	69	69	26
4	0.999	110	69	27
4	0.998	105	65	30
4	0.994	66	147	109
3	0.996	72	75	47
3	0.993	82	67	57
2	0.995	63	84	102
2	0.991	108	139	130
2	1.000	116	86	108
2	0.991	46	141	97
2	0.991	99	106	61
2	0.997	77	149	64
2	0.992	76	146	66
1	0.991	53	142	66
1	0.992	67	84	105
1	0.991	49	141	97
1	0.991	77	81	52
1	0.990	77	72	51
1	0.991	74	71	50
1	0.990	103	63	49
1	0.990	77	70	45
1	0.992	107	57	42
1	0.990	107	49	42
1	0.990	80	89	33
1	0.990	72	68	32
1	0.990	108	77	25
1	0.990	103	61	50
PRS in REL >PRS in SZ				
27	0.995	67	82	45
11	0.993	75	72	50
5	0.999	103	119	80
5	0.998	110	174	97
4	0.993	108	69	27
3	0.991	65	68	43
3	0.991	62	65	46
2	0.991	62	65	43
2	0.991	74	76	48
2	0.991	105	116	83
1	0.991	61	68	45
1	0.992	72	75	47
1	0.992	48	145	83

TABLE B.6: (Continuation) TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in BD >PRS in REL				
7	0.999	108	76	46
2	0.996	105	65	30
2	0.998	63	83	102
2	0.994	61	107	103
2	0.998	116	86	108
1	0.991	60	108	101
1	0.993	109	161	110
1	0.995	132	88	112

TABLE B.7: TFCE-uncorrected PRS x diagnosis interaction effects on FA with less than 50 voxels, for the age-matched sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on FA				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in SZ >PRS in HC				
3	0.993	46	75	106
2	0.991	107	152	123
1	0.992	126	155	94
1	0.99	81	114	127
PRS in BD >PRS in HC				
5	0.993	67	82	45
5	0.995	69	85	45
1	0.99	66	80	45
1	0.99	68	85	47
1	0.993	118	128	90
1	0.99	135	108	94
PRS in REL >PRS in HC				
1	0.992	109	125	74
1	0.992	53	169	78
1	0.991	67	179	85
PRS in HC >PRS in SZ				
2	0.999	100	67	51
1	0.993	66	71	27
PRS in HC >PRS in REL				
3	0.998	98	108	128
2	0.997	100	67	51
1	0.993	108	74	47
1	0.991	110	162	109
PRS in BD >PRS in SZ				
2	0.997	100	67	50
PRS in SZ >PRS in BD				
2	0.999	59	175	75
1	0.993	126	155	94
PRS in REL >PRS in SZ				
1	0.994	66	71	27
PRS in SZ >PRS in REL				
3	0.995	98	108	127
2	0.998	62	74	97
1	0.992	126	155	94
PRS in BD >PRS in REL				
5	0.999	98	108	128
2	0.992	106	77	48
1	0.992	100	67	50
1	0.99	124	51	85
1	0.991	42	82	111

TABLE B.8: TFCE-uncorrected PRS x diagnosis interaction effects on MD with less than 50 voxels, for the whole-sample. Each effect is characterised by the cluster extent,1-(p-value) and the MNI peak coordinates.

PRS x diagnosis effects on MD				
Cluster extent (k)	**1-(p-value)**	**Peak MNI coordinates**		
		x{mm}	**y{mm}**	**z{mm}**
PRS in HC >PRS in SZ				
2	0.992	90	70	54
2	0.993	120	126	54
2	0.992	52	166	82
1	0.992	119	129	56
1	0.994	51	52	69
1	0.993	98	163	70
PRS in HC >PRS in BD				
9	0.998	99	113	127
6	0.994	127	140	68
3	0.997	99	108	66
2	0.992	74	77	102
1	0.995	105	65	30
1	0.992	123	140	68
1	0.995	112	99	69
1	0.996	55	150	99
1	0.992	74	80	102
1	0.992	66	85	116
PRS in HC >PRS in REL				
1	0.992	80	41	89
1	0.992	79	42	91
PRS in BD >PRS in SZ				
1	0.990	90	69	55
PRS in SZ >PRS in BD				
2	0.995	105	65	30
2	0.992	99	113	127
1	0.990	106	68	29
1	0.990	100	110	63
PRS in REL >PRS in SZ				
5	0.992	92	71	50
5	0.996	89	69	57
1	0.994	90	72	47
PRS in SZ >PRS in REL				
2	0.992	100	110	63
1	0.992	98	109	62
PRS in BD >PRS in REL				
1	0.994	70	184	73
PRS in REL >PRS in BD				
1	0.994	105	65	30
1	0.995	103	135	64
1	0.991	99	113	127